Vol. 22. No 1, 2020

ISSN 1507-2711
Cena: 25 zł (w tym 8% VAT)

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY

Polskie Naukowo Techniczne Towarzystwo Eksploatacyjne
Warszawa

Polish Maintenance Society
Warsaw

KILIKEVIČIUS A, RIMŠA V, RUCKI M. **Investigation of influence of aircraft propeller modal parameters on small airplane performance**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 1–5, http://dx.doi.org/10.17531/ein.2020.1.1.

The aim of the current paper is to investigate a small airplane model propeller of class F2D according to requirements of Fédération Aéronautique Internationale (FAI, or World Air Sports Federation). In some cases, practical tests show that F2D models with flexible propellers produce specific extra noise and increase flight speed in comparison with "rigid" propellers. Therefore, the following hypothesis could be proposed: flexible characteristics of the increased noise are related to the resonant eigenfrequencies of the propeller. The operating range of the F2D class propeller (28,000-35,000 rpm) is close to or equal to the eigenfrequency resonance. The current investigation addresses dynamic/flexible vibrations of elastic propeller during engine run and researches dynamic parameters of the propeller as well as the contribution of these parameters to the model flight characteristics. To resolve this type of a problem, a stand, which allows completing a physical investigation of flexible propeller vibration modes and dynamic characteristics was created.

KILIKEVIČIUS A, RIMŠA V, RUCKI M. **Badania wpływu modalnych parametrów śmigła na zachowanie małego samolotu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 1–5, http://dx.doi.org/10.17531/ein.2020.1.1.

Celem artykułu jest przedstawienie wyników badań śmigła małego modelu samolotu zaliczanego do klasy F2D (według klasyfikacji Fédération Aéronautique Internationale, FAI). W niektórych przypadkach testy wykazały, że modele F2 z giętkimi śmigłami, w porównaniu do śmigieł sztywnych, wydają dodatkowy hałas i zwiększają prędkość samolotu. Dlatego wysunięto hipotezę, że elastyczne charakterystyki zwiększonego hałasu są powiązane z rezonansem częstotliwości własnych śmigła. Zakres pracy śmigła klasy F2D (28 000-35 000 obr/min) jest zbliżony do jego częstotliwości własnych. Badania dotyczą elastycznych wibracji dynamicznych śmigła giętkiego w czasie rozruchu silnika i są nakierowane na wyznaczanie parametrów dynamicznych i ich wpływu na charakterystyki lotu modelu. Wykonano i opisano stanowisko, na którym przeprowadzono testy modalne drgań giętkiego śmigła. Na tej podstawie uzyskano charakterystyki dynamiczne.

HUANG H-Z, YU K, HUANG T, Li H, QIAN H-M. **Reliability estimation for momentum wheel bearings considering frictional heat**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 6–14, http://dx.doi.org/10.17531/ein.2020.1.2.

Momentum wheels are the key components of the inertial actuators in the satellites, and the momentum wheel bearings are weak links of momentum wheels as they operate under harsh conditions. The reliability estimation for momentum wheel bearings are helpful to guarantee the mission successes for both momentum wheels and satellites. Hence, this paper put emphasis into reliability estimation of a momentum wheel bearing considering multiple coupling operating conditions and frictional heat by using the finite element analysis. The stress-strength interference model is employed to calculate the reliability of the momentum wheel bearing. A comparative analysis for reliability estimation with and without frictional heat of the momentum wheel bearing is conducted. The results show that the frictional heat cannot be ignored in the reliability analysis of momentum wheel bearings.

HUANG H-Z, YU K, HUANG T, Li H, QIAN H-M. **Ocena niezawodności łożysk kół zamachowych z uwzględnieniem ciepła tarcia**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 6–14, http://dx.doi.org/10.17531/ein.2020.1.2.

Koła zamachowe są kluczowymi elementami składowymi siłowników bezwładnościowych w satelitach. Ich łożyska stanowią słaby punkt podczas pracy w trudnych warunkach. Ocena niezawodności łożysk kół zamachowych pozwala zapewnić powodzenie misji zarówno w odniesieniu do samych kół zamachowych, jak i satelitów. Dlatego też niniejszy artykuł poświęcono zagadnieniu oceny niezawodności łożyska koła zamachowego z wykorzystaniem analizy metodą elementów skończonych przy uwzględnieniu wielu sprzężonych warunków pracy oraz ciepła tarcia. Do obliczenia niezawodności łożyska koła zamachowego zastosowano model obciążeniowo-wytrzymałościowy. Przeprowadzono także analizę porównawczą oceny niezawodności łożyska koła zamachowego z uwzględnieniem lub bez uwzględnienia ciepła tarcia. Wyniki pokazują, że w analizie niezawodności łożysk kół zamachowych nie można pominąć ciepła tarcia.

ANDRUSZKIEWICZ J, LORENC J, ŁOWCZOWSKI K, WEYCHAN A, ZAWODNIAK J. **Energy losses' reduction in metallic screens of MV cable power lines and busbar bridges composed of single-core cables**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 15–25, http://dx.doi.org/10.17531/ein.2020.1.3.

The growing share of medium voltage cable lines in distribution networks challenges distribution network operators in terms of proper mode of operation of these lines. It is related to the reduction of energy losses in cable conductors and metallic cable screens. The article focuses on energy losses in metallic cable screens of cable lines and substation busbar bridges composed of single-core cables with metallic screens and possible ways of their reduction. Simulation and measurement analysis of the level of energy losses in the metallic screens of cables is presented together with the economic analysis of various variants of losses reduction through the change of the way these screens are operated in relation to the traditional bilateral earthing at both ends of cable. Technical problems and threats connected with the use of considered modifications of metallic screens operation during earth fault disturbances in distribution networks are also presented.

ANDRUSZKIEWICZ J, LORENC J, ŁOWCZOWSKI K, WEYCHAN A, ZAWODNIAK J. **Ograniczanie strat energii w żyłach powrotnych linii i mostów kablowych średniego napięcia wykorzystujących kable jednożyłowe**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 15–25, http://dx.doi.org/10.17531/ein.2020.1.3.

Rosnące skablowanie linii średniego napięcia w sieciach dystrybucyjnych stawia przed operatorami tych sieci wyzwanie prawidłowej eksploatacji linii kablowych. Powiązane jest to z redukowaniem strat energii w żyłach roboczych i powrotnych kabli. W artykule skupiono się na stratach energii w żyłach powrotnych linii oraz mostów kablowych wykonanych przy wykorzystaniu kabli jednożyłowych z metalicznymi żyłami powrotnymi oraz możliwych sposobach ich ograniczania. Przedstawiono analizę symulacyjną i pomiarową poziomu strat energii w żyłach powrotnych kabli wraz z analizą ekonomiczną różnych wariantów ich redukcji poprzez zmianę sposobu pracy tych żył w stosunku do tradycyjnego obustronnego ich uziemienia. Przedstawione zostały również problemy techniczne oraz zagrożenia związane z zastosowaniem rozważanych modyfikacji pracy żył powrotnych podczas zakłóceń zwarciowych w sieciach dystrybucyjnych.

ZHANG C, ZHANG Y. **Common cause and load-sharing failures-based reliability analysis for parallel systems**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 26–34, http://dx.doi.org/10.17531/ein.2020.1.4.

For parallel system reliability, the mean time to failure of parallel system under common cause failure (load-sharing failure) is shorter than that of the system without common cause failure (load-sharing failure). The traditional calculation approaches of mean time to failure of parallel systems do not consider the possible effect of common cause and load-sharing failure. However, it may result in the poor accuracy of mean time to failure of parallel system and pose a threat to system reliability. This paper not only considers the effect of common cause failure with stress strength, but also investigates the joint effect of the load-sharing and common cause failures. Besides, the joint failure model of three-dependent-component parallel system are established, and the corresponding properties are analyzed. Finally, a numerical example is used to illustrate the proposed method.

ZHANG C, ZHANG Y. **Analiza niezawodności systemów równoległych w sytuacji jednocześnie występujących uszkodzeń wywołanych wspólną przyczyną oraz uszkodzeń elementów dzielących obciążenie**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 26–34, http://dx.doi.org/10.17531/ein.2020.1.4.

Gdy mowa o niezawodności systemu równoległego, średni czas do uszkodzenia, w przypadku uszkodzenia wywołanego wspólną przyczyną (lub uszkodzenia elementów dzielących obciążenie) jest krótszy niż dla systemu, w którym nie występują tego typu uszkodzenia. Tradycyjne metody obliczania średniego czasu do uszkodzenia systemów równoległych nie uwzględniają potencjalnego wpływu uszkodzeń wywołanych wspólną przyczyną oraz uszkodzeń komponentów dzielących obciążenie. Może to skutkować małą dokładnością tak obliczanego średniego czasu do uszkodzenia systemu równoległego i stanowić zagrożenie dla jego niezawodności. W prezentowanej pracy rozważano nie tylko wpływ uszkodzenia wywołanego wspólną przyczyną dla modelu typu wytrzymałość-obciążenie, ale również wpływ jednocześnie występujących uszkodzeń wywołanych wspólną przyczyną i uszkodzeń elementów dzielących obciążenie. Poza tym opracowano model, w którym omawiane dwa typy uszkodzeń występują jednocześnie w systemie równoległym składającym się z trzech zależnych elementów oraz przeanalizowano właściwości takiego systemu. W artykule przedstawiono przykład numeryczny, który ilustruje zastosowanie proponowanej metody.

MŁYNARSKI S, PILCH R, SMOLNIK M, SZYBKA J, WIĄZANIA G. **A model of an adaptive strategy of preventive maintenance of complex technical objects**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 35–41, http://dx.doi.org/10.17531/ein.2020.1.5.

The paper presents results of the analysis of the developed models for complex technical objects preventive maintenance scheduling. Models based on two different sets of assumptions were developed. The general problem solved was to determine the joint time of preventive renewal for a group of parts or subassemblies. The purpose of the first model (the model of scheduled preventive maintenance strategy) is to determine the profitability of constant application of a previously developed preventive maintenance schedule for a part undergoing post-failure renewal. The second model (the model of adaptive strategy of a system's preventive maintenance) allows one to determine a new joint time of preventive renewal for a group of parts each time when one of them is undergoing a post-failure renewal. The initial preventive maintenance strategy for each part or subassembly was obtained using typical tools for maintenance planning (decision-random models based on dynamic programming and Bellman's principle of optimality). Exemplary simulation calculations with the use of both models were made and their results presented as the total maintenance costs estimated for the renewal strategies developed. The object of the analysis were the chosen geometrical features of a rail vehicle wheel changing due to its wear during operation. Based on this kind of analysis, one can choose a better preventive maintenance model for a specific application area.

MŁYNARSKI S, PILCH R, SMOLNIK M, SZYBKA J, WIĄZANIA G. **Model adaptacyjnej strategii prewencyjnej odnowy złożonych obiektów technicznych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 35–41, http://dx.doi.org/10.17531/ein.2020.1.5.

W artykule przedstawiono wyniki analizy opracowanych modeli do planowania odnowy profilaktycznej złożonych obiektów technicznych, które oparto o dwa różniące się od siebie zestawy założeń. Rozwiązywany problem dotyczy określenia wspólnego czasu odnowy profilaktycznej grupy części lub podzespołów złożonego obiektu. Pierwszy z opracowanych modeli (model planowej strategii odnowy prewencyjnej) pozwala określić zasadność przeprowadzenia ustalonego wcześniej, planowego odnowienia prewencyjnego części obiektu, która została już odnowiona poawaryjnie. Drugi model (model adaptacyjnej strategii odnowy prewencyjnej) umożliwia wyznaczenie najbliższego wspólnego czasu odnowy profilaktycznej grupy części, z których jedna aktualnie podlega odnowie poawaryjnej. Początkowe (wyjściowe) strategie odnowy profilaktycznej każdej części bądź podzespołu wyznaczone zostały za pomocą standardowych narzędzi do planowania odnawiania profilaktycznego (modeli decyzyjno-losowych wykorzystujących programowanie dynamiczne Bellmana). Posługując się opracowanymi modelami odnowy, przeprowadzono przykładowe obliczenia symulacyjne, których wyniki przedstawiono w postaci całkowitych kosztów obsługiwania dla każdej z uzyskanych strategii. Przedmiotem analizy były wybrane cechy geometryczne koła pojazdu szynowego, których wartości zmieniają się na skutek zużycia w procesie eksploatacji. Na podstawie tego rodzaju analiz można wybrać lepszy (tj. efektywniejszy ekonomicznie) z modeli dla konkretnego zastosowania w praktyce.

CUI J, REN Y, XU B, YANG D, ZENG S. **Reliability analysis of a multi-eso based control strategy for level adjustment control system of quadruped robot under disturbances and failures**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 42–51, http://dx.doi.org/10.17531/ein.2020.1.6.

The complexity of control algorithms and their vulnerability to disturbances and failures are the main problems that restrict the operations of multi-legged mobile robots in more complex environments. In this paper, a multiple extended state observer (ESO) based control strategy is proposed to achieve stable tilt angle control for quadruped robots under the influence of disturbances and actuator failures. By treating the multiple legs as parallel control objects, more ESOs were added to improve the disturbance rejection ability of the linear active disturbance rejection control (LADRC). Correlation of interactive information about the legs is realized by the synthesis of multiple ESO information. Based on LADRC, this method has the advantages of easy parameter tuning, good robustness, and strong ability to cope with interference and fault conditions. A control system reliability evaluation method was proposed. The reliability and control performance of the multi-ESO based control system under leg stuck failure conditions were systematically analyzed. Simulation and experimental results for the level adjustment control system of a quadruped robot are provided to verify the disturbance rejection ability, feasibility and practicability of the proposed multi-ESO based control method.

CUI J, REN Y, XU B, YANG D, ZENG S. **Analiza niezawodności strategii sterowania w oparciu o kilka obserwatorów stanu rozszerzonego ESO i jej zastosowanie w systemie kontroli regulacji poziomu czworonożnego robota w warunkach zakłóceń i uszkodzeń**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 42–51, http://dx.doi.org/10.17531/ein.2020.1.6.

Złożoność algorytmów sterowania oraz ich brak odporności na zakłócenia i uszkodzenia to główne czynniki ograniczające pracę wielonożnych robotów mobilnych w bardziej złożonych środowiskach. W przedstawionym artykule zaproponowano strategię sterowania wykorzystującą kilka obserwatorów stanu rozszerzonego (ESO), która pozwala na uzyskanie stabilnego kąta pochylenia robota czworonożnego w warunkach zakłóceń i uszkodzeń siłowników. Traktując każdą z nóg jako równorzędny obiekt sterowania, dodano dodatkowe ESO, co pozwoliło na poprawienie zdolności algorytmu liniowego aktywnego tłumienia zakłóceń (LADRC) do kompensacji (tłumienia) tych ostatnich. Interaktywne informacje dotyczące poszczególnych nóg korelowano poprzez syntezę danych z ESO. Zaletami omawianej metody opartej na LADRC są: łatwość dostrajania parametrów, wysoka niezawodność oraz bardzo dobra zdolność do radzenia sobie z zakłóceniami i uszkodzeniami. Zaproponowano także metodę oceny niezawodności systemu sterowania. Analizowano niezawodność i wydajność systemu opartego na kilku ESO w warunkach awarii wywołanej zablokowaniem nóg robota. Przedstawiono wyniki badań symulacyjnych i eksperymentalnych systemu sterowania regulacją poziomu robota czworonożnego, które pozwalają zweryfikować zdolność proponowanej metody do tłumienia zakłóceń, a także możliwość jej praktycznego zastosowania.

SOBASZEK Ł, GOLA A, ŚWIĆ A. **Time-based machine failure prediction in multi-machine manufacturing systems**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 52–62 http://dx.doi.org/10.17531/ein.2020.1.7.

The execution of production processes in real manufacturing systems is associated with the occurrence of numerous disruptions, which predominantly revolve around technological machine failure. Therefore, various maintenance strategies are being developed, many of which tend to emphasise effective preventive measures, such as the Time-Based Maintenance (TBM) discussed in this paper. Specifically, this publication presents the time-based machine failure prediction algorithm for the multi-machine manufacturing environment. The Introduction section outlines the body of knowledge related to typical strategies applied in maintenance. The next part describes an approach to failure prediction that treats processing times as makespan and is followed by highlighting the key role of historical data in machine failure management, in the subsequent section. Finally, the proposed time-based machine failure prediction algorithm is presented and tested by means of a two-step verification, which confirms its effectiveness and further practical implementation.

SOBASZEK Ł, GOLA A, ŚWIĆ A. **Algorytm wsparcia strategii TBM w wielomaszynowych systemach wytwórczych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 52–62 http://dx.doi.org/10.17531/ein.2020.1.7.

Realizacja procesów produkcyjnych w rzeczywistych systemach wytwórczych wiąże się z występowaniem wielu zakłóceń, do których zalicza się głównie awarie maszyn technologicznych. W związku z tym obserwowany jest rozwój różnorodnych strategii utrzymania ruchu. Coraz większy nacisk kładziony jest na efektywne działania prewencyjne, do których zalicza się także działania określone w czasie (ang. Time-Based Maintenance – TBM). W niniejszej publikacji zaprezentowano algorytm predykcji awarii maszyn w wielomaszynowych systemach wytwórczych wspierający prewencyjne utrzymanie ruchu. Na wstępie omówiono zagadnienia związane z typowymi strategiami stosowanymi w obszarze UR. Ponadto omówiono tematykę predykcji awarii, zwracając uwagę na ujęcie czasu pracy maszyny jako czasu trwania, a także kluczową rolę wykorzystania danych historycznych dotyczących awarii maszyn. Następnie zaprezentowano proponowany algorytm predykcji wspierający działania określone w czasie. Prezentowane prace zakończono dwuetapową weryfikacją proponowanej metody, która potwierdziła jej skuteczność oraz zasadność wykorzystania.

LI Y, WANG K. **Modified convolutional neural network with global average pooling for intelligent fault diagnosis of industrial gearbox**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 63–72, http://dx.doi.org/10.17531/ein.2020.1.8.

Gearboxes are key transmission components and widely used in various industrial applications. Due to the possible operational conditions, such as varying rotational speeds, long period of heavy loads, etc., gearboxes may easily be prone to failure. Condition Monitoring (CM) has been proved to be an effective methodology to improve the safety and reliability of gearboxes. Deep learning approaches, nowadays, further enable the CM with more powerful capability to exploit faulty information

LI Y, WANG K. **Diagnostyka błędów przekładni przemysłowych z wykorzystaniem zmodyfikowanej splotowej sieci neuronowej z globalnym uśrednieniem wartości dla poszczególnych kanałów**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 63–72, http://dx.doi.org/10.17531/ein.2020.1.8.

Przekładnie stanowią kluczowe elementy układów napędowych i jako takie znajdują szerokie zastosowane w przemyśle. Ze względu na warunki eksploatacji, takie jak różne prędkości obrotowe czy długie okresy pracy pod dużym obciążeniem itp., przekładnie mogą łatwo ulegać uszkodzeniom. Udowodniono, że monitorowanie stanu skutecznie poprawia bezpieczeństwo i niezawodność przekładni. Podejścia oparte na uczeniu głębo-

from massive data and make intelligently diagnostic decisions. However, for most of conventional deep learning models, such as Convolutional Neural Network (CNN), a large amount of labelled training data is a prerequisite, while to obtain the labelled data is usually a laborious and time-consuming job and sometimes even unattainable. In this paper, to handle the case of only a limited labelled data is available, a modified convolutional neural network (MCNN) is proposed by integrating global average pooling (GAP) to reduce the number of trainable parameters and simplify the architecture of deep learning model. The proposed MCNN improves the traditional CNN's ability in fault diagnosis with limited labelled data. Two experimental gearbox datasets are utilized to demonstrate the effectiveness of the proposed MCNN method. Compared with traditional deep learning approaches, namely LSTM, CNN and its variant methods, the experimental results show that the proposed MCNN with higher discrimination and generalization ability in fault classification and diagnostics under the scenario of limited labelled training samples.

kim umożliwiają ponadto monitorowanie stanu z większym wykorzystaniem informacji o błędach pochodzących z dużych zbiorów danych i podejmowanie inteligentnych decyzji diagnostycznych. Jednak w przypadku większości konwencjonalnych modeli uczenia głębokiego, takich jak splotowe sieci neuronowe (convolutional neural networks, CNN), wymagana jest duża ilość etykietowanych danych uczących, których pozyskanie jest zwykle zadaniem praco- i czasochłonnym, a czasem wręcz niemożliwym do wykonania. W niniejszej pracy, przedstawiono zmodyfikowaną splotową sieć neuronową (modified convolutional neural network, MCNN), która rozwiązuje problem dostępności danych etykietowanych poprzez zastosowanie globalnego uśrednienia względem kanałów (global average pooling), co pozwala na zmniejszenie liczby możliwych do wyuczenia parametrów i uproszczenie architektury modelu głębokiego uczenia. W porównaniu do tradycyjnych sieci CNN, proponowana sieć MCNN zwiększa możliwości diagnozowania błędów przy ograniczonych danych etykietowanych. Skuteczność proponowanej metody wykazano na przykładzie dwóch zbiorów danych doświadczalnych dotyczących błędów przekładni. Wyniki eksperymentalne pokazują, że, w porównaniu z tradycyjnymi metodami uczenia głębokiego, takimi jak LSTM, CNN oraz warianty tej ostatniej, proponowane podejście MCNN daje większe możliwości rozróżniania i uogólniania podczas klasyfikacji i diagnostyki błędów w przypadku ograniczonej dostępności etykietowanych danych uczących.

KONOWROCKI R, CHOJNACKI A. **Analysis of rail vehicles' operational reliability in the aspect of safety against derailment based on various methods of determining the assessment criterion**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 73–85, http://dx.doi.org/10.17531/ein.2020.1.9.

The article features the results of computer and experimental research on operational issues in the aspect of safety in relation to a freight wagon derailment on a railway track. It presents the knowledge regarding the methods of assessing the operational safety of rail vehicles on railroad tracks for the purpose of comparative analysis. The theoretical analyses were performed based on several methods that assess the safety of their derailments, qualifying for operational reliability, comparing them with the results obtained from experimental research. For the purpose of the research, a computer model of rail vehicle- railway track was created. It took into consideration dynamic parameters of elements used in the real track and rail vehicle. The results obtained from theoretical analyses were validated with experimental tests carried out on real objects (freight vehicle - test track, freight wagon - test rig). As part of the research, new test track geometry for testing rail vehicles was proposed. The results obtained in this way allowed estimating the conditions threatening the operation of a freight vehicle while running on the test rail infrastructure with different assessment criteria and to compare them.

KONOWROCKI R, CHOJNACKI A. **Analiza niezawodności eksploatacyjnej pojazdów szynowych w aspekcie bezpieczeństwa przed wykolejeniem w oparciu o różne metody wyznaczania kryterium oceny**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 73–85, http://dx.doi.org/10.17531/ein.2020.1.9.

W pracy pokazano rezultaty badań komputerowych i eksperymentalnych dotyczących zagadnień eksploatacji w aspekcie bezpieczeństwa w odniesieniu do wykolejenia wagonu towarowego na torze kolejowym. Przybliżono w nim stan wiedzy dotyczącej metod oceny bezpieczeństwa eksploatacji pojazdówa) szynowych na kolejowych liniach szynowych, w celu ich analizy porównawczej. W pracy wykonano analizy teoretyczne bazując na kilku metodach, które oceniają bezpieczeństwo ich wykolejenia, kwalifikujące się do niezawodność eksploatacyjnej, porównując je z wynikami otrzymanymi z badań eksperymentalnych. Na potrzeby przeprowadzanych badań powstał komputerowy model pojazd szynowy - tor kolejowy. Uwzględniał on parametry dynamiczne elementów zastosowanych w rzeczywistym torze oraz pojeździe szynowym. Otrzymane z teoretycznych analiz wyniki zwalidowano testami eksperymentalnymi wykonanymi na rzeczywistych obiektach (pojazd towarowy - tor testowy, wagon towarowy - stanowisko badawcze). W ramach badań zaproponowano nową geometrię toru testowego do badań pojazdów szynowych. Uzyskane wyniki pozwoliły określić stan zagrożenia eksploatacji wagonu towarowego podczas jazdy po testowej infrastrukturze szynowej przy różnych kryteriach oceny oraz je porównać.

WANG Z-B, LI W-Y, SHANG S, WANG Z, HAN C-Y. **Performance degradation comparisons and failure mechanism of silver metal oxide contact materials in relays application by simulation**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 86–93, http://dx.doi.org/10.17531/ein.2020.1.10.

To evaluate the electrical contact behaviors of silver metal oxide contact materials in relays application more accurately, and to guide the selection of contact materials, the test device and testing method for simulating electrical contact performance in relays application were analyzed in this paper. The electrical contact simulation test system was designed and developed, which can easily simulate contact materials. The contact resistance, static force and rebound energy degradation parameters of $AgSnO_2$, AgCdO and AgNi contact materials under the same load conditions were obtained through experimental research, the contact resistance and arcing energy degradation parameters of $AgSnO_2$ under different opening distances were acquired at the same time. The result indicated that accurate data are received by the electrical contact simulation testing method. Finally, based on the test data, the degradation performance of three selected test materials was tested, and the failure mechanism of $AgSnO_2$ materials was analyzed.

WANG Z-B, LI W-Y, SHANG S, WANG Z, HAN C-Y. **Porównanie obniżenia charakterystyk oraz badanie mechanizmu uszkodzeń materiałów stykowych z kompozytów typu srebro-tlenek metalu stosowanych w przekaźnikach elektromagnetycznych na podstawie danych z symulacji komputerowej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 86–93, http://dx.doi.org/10.17531/ein.2020.1.10.

W celu dokładniejszej oceny zachowania styków elektrycznych z kompozytów srebra i tlenku metalu stosowanych w przekaźnikach elektromagnetycznych oraz w celu ułatwienia wyboru materiałów stykowych, w niniejszej pracy przeanalizowano urządzenie testowe oraz metodę testowania, które pozwalają na symulację działania styku przekaźnika. Zaprojektowano i zbudowano system testowania styków elektrycznych, który umożliwia łatwą symulację zachowania materiałów stykowych. Parametry degradacji rezystancji zestykowej, siły statycznej oraz energii odbicia materiałów stykowych $AgSnO_2$, AgCdO i AgNi uzyskano w badaniach eksperymentalnych prowadzonych w takich samych warunkach obciążenia. Jednocześnie badano także parametry degradacji rezystancji zestykowej energii łuku $AgSnO_2$ przy różnych odległościach otwarcia styków. Wyniki pokazują, że proponowana metoda badania symulacyjnego styków elektrycznych pozwala na uzyskanie dokładnych danych. W oparciu o dane testowe, przebadano zachowanie degradacyjne trzech wybranych materiałów oraz przeanalizowano mechanizm uszkodzenia styków z kompozytu $AgSnO_2$.

KNEFEL T, NOWAKOWSKI J. **Model-based analysis of injection process parameters in a common rail fuel supply system**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 94–101, http://dx.doi.org/10.17531/ein.2020.1.11.

In this work, a simplified model of a rail-based fuel supply system in a compression ignition engine is presented. High pressure hoses were not taken into consideration and an empirical model was developed to simulate the injectors. The basic equations of the model are presented. Phenomena were described by 17 first-order ordinary differential equations. This work also contains an evaluation of the impact of the rail's geometrical parameters on the injection process. The evaluation was carried out via a program calculating the injection process, using a balance model of the injection system. A method for making preliminary decisions on the geometrical parameters of the rail is proposed.

KNEFEL T, NOWAKOWSKI J. **Modelowa analiza parametrów procesu wtrysku w układzie zasilania typu common rail**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 94–101, http://dx.doi.org/10.17531/ein.2020.1.11.

W pracy przedstawiono uproszczony model zasobnikowego układu zasilania w paliwo silnika o zapłonie samoczynnym. W rozważaniach nie uwzględniono przewodów wysokiego ciśnienia, a do symulacji pracy wtryskiwaczy opracowano empiryczny podmodel. Przedstawiono podstawowe równania modelu. Zjawiska zostały opisane układem 17 równań różniczkowych zwyczajnych, pierwszego rzędu. W pracy również zawarto ocenę wpływu parametrów geometrycznych zasobnika na proces wtrysku. Ocenę przeprowadzono za pomocą programu obliczającego proces wtrysku, wykorzystującego model rozważanego układu wtryskowego. Zaproponowano sposób wstępnego doboru parametrów geometrycznych zasobnika.

SELECH J, ANDRZEJCZAK K. **An aggregate criterion for selecting a distribution for times to failure of components of rail vehicles**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 102–111, http://dx.doi.org/10.17531/ein.2020.1.12.

Abstract This paper presents an aggregate method of selecting a theoretical cumulative distribution function (CDF) for an empirical CDF. The method was intended to identify the time of reliable operation of a renewable technical object by applying three criteria based on the following statistics: the modified Kolmogorov–Smirnov (MK-S) statistic, the mean absolute deviation of the theoretical CDF from the empirical CDF, and a statistic calculated on the basis of a log-likelihood function. The values of these statistics were used to rank eleven probability distributions. The data for which calculations were made concerned failures of the driver's cab lock recorded during five years of operation of a fleet of 45 trams. Before calculating the statistics, the empirical CDF of the examined component was determined using the Kaplan–Meier estimator, and then, using the method of Maximum Likelihood Estimation, the parameters of the analysed theoretical distributions were estimated. The theoretical distributions were then ranked according to the values obtained for each of the assumed criteria: the lower the value for a given criterion, the higher the ranking position, indicating a better fit according to that criterion. Then, based on the three rankings and on weights assigned to the individual criteria, an aggregate criterion (referred to as DESV) was implemented to select the best-fitting probability distribution. The method assumes that the lowest DESV value corresponds to the best-fitting theoretical distribution. In the case of the examined component, this was found to be the generalised gamma distribution. It is shown that if the final decision is based on the aggregate criterion, which takes into account the three criteria for goodness of fit, the reliability of the estimation of the time-to-failure distribution increases, and thus mistakes resulting from the use of only one of the criteria can be avoided.

SELECH J, ANDRZEJCZAK K. **Zagregowane kryterium wyboru rozkładu czasu do uszkodzenia elementów pojazdów szynowych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 102–111, http://dx.doi.org/10.17531/ein.2020.1.12.

W pracy przedstawiono zagregowaną metodę doboru dystrybuant hipotetycznych do dystrybuanty empirycznej. Metoda miała na celu identyfikację czasu niezawodnej pracy odnawialnego obiektu technicznego poprzez zastosowanie trzech kryteriów, w których użyto następujących statystyk: zmodyfikowanej statystyki Kołmogorowa-Smirnowa (MK-S), statystyki średniego odchylenia bezwzględnego dystrybuanty hipotetycznej od empirycznej oraz statystyki obliczanej na podstawie zlogarytmowanej funkcji wiarygodności. Wartości tych statystyk posłużyły do rangowania jedenastu rozkładów prawdopodobieństwa. Dane dla których dokonano obliczeń dotyczyły uszkodzeń zamka kabiny motorniczego jakie odnotowano w ciągu pięciu lat użytkowania floty 45 tramwajów. Przed obliczeniem statystyk wyznaczono dystrybuantę empiryczną badanego elementu przy pomocy estymatora Kaplana-Meiera, a następnie przy użyciu metody największej wiarygodności oszacowano parametry uwzględnionych w badaniach rozkładów hipotetycznych. Po wyznaczaniu parametrów nastąpiło rangowanie rozkładów hipotetycznych według wartości otrzymanych dla każdego z przyjętych kryteriów, im mniejsza wartość dla danego kryterium tym wyższa pozycja w rankingu, świadcząca o lepszej jakości dopasowania według danego kryterium. Po ustaleniu rankingu według kryteriów zgodności, każdemu z kryteriów zgodności dopasowania dystrybuant modelowych do empirycznej nadano wagi. Następnie na podstawie uzyskanych trzech rankingów oraz wag nadanych poszczególnym kryteriom zgodności wyznaczana jest zagregowana miara zgodności (oznaczona DESV), która służy do wyznaczania najlepszego rozkładu prawdopodobieństwa. W prezentowanej metodzie przyjęto, że najmniejsza wartość DESV wyznacza najlepiej dopasowany rozkład hipotetyczny. W przypadku badanego elementu rozkładem tym okazał się uogólniony rozkład gamma. Pokazano, że na podstawie zagregowanego kryterium uwzględniającego trzy statystyki zgodności dopasowania zwiększa się wiarygodność estymacji rozkładu czasu pracy do uszkodzenia, unikając tym samym błędów jakie można popełnić uzależniając się tylko od jednej z nich.

MI J, LI Y-F, BEER M, BROGGI M, CHENG Y. **Importance measure of probabilistic common cause failures under system hybrid uncertainty based on bayesian network**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 112–120, http://dx.doi.org/10.17531/ein.2020.1.13.

When dealing with modern complex systems, the relationship existing between components can lead to the appearance of various dependencies between component failures, where multiple items of the system fail simultaneously in unpredictable fashions. These probabilistic common cause failures affect greatly the performance of these critical systems. In this paper a novel methodology is developed to quantify the importance of common cause failures when hybrid uncertainties are presented in systems. First, the probabilistic common cause failures are modeled with Bayesian networks and are incorporated into the system exploiting the α factor model. Then, probability-boxes (bound analysis method) are introduced to model the hybrid uncertainties and quantify the effect of uncertainties on system reliability. Furthermore, an extended Birnbaum importance measure is defined to identify the critical common cause failure events and coupling impact factors when uncertainties are expressed by probability-boxes. Finally, the effectiveness of the method is demonstrated through a numerical example.

MI J, LI Y-F, BEER M, BROGGI M, CHENG Y. **Oparta na sieci bayesowskiej miara ważności probabilistycznych uszkodzeń spowodowanych wspólną przyczyną w warunkach niepewności hybrydowej systemu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 112–120, http://dx.doi.org/10.17531/ein.2020.1.13.

W przypadku nowoczesnych systemów złożonych, relacje zachodzące między komponentami mogą prowadzić do pojawienia się różnych zależności między ich uszkodzeniami, a tym samym do sytuacji w których kilka składowych systemu ulega uszkodzeniu jednocześnie w nieprzewidywalny sposób. Tego typu probabilistyczne uszkodzenia wywołane wspólną przyczyną (PCCF) mają ogromny wpływ na wydajność tych kluczowych systemów. W przedstawionym artykule opracowano nową metodę szacowania ważności PCFF w sytuacjach, gdy w systemie występują niepewności hybrydowe. W pierwszej kolejności, PCFF zamodelowano za pomocą sieci bayesowskich i włączono do systemu wykorzystującego model współczynnika α. Następnie, wprowadzono przedziały prawdopodobieństwa, tzw. probability boxes (bound analysis method), w celu zamodelowania niepewności hybrydowych i kwantyfikacji wpływu tych niepewności na niezawodność systemu. Ponadto zdefiniowano rozszerzoną miarę ważności Birnbauma, która pozwala zidentyfikować krytyczne zdarzenia PCCF oraz czynniki, które je wywołały, w przypadkach, gdy niepewności wyrażone są za pomocą probability boxes. Skuteczność metody wykazano na przykładzie numerycznym.

JÓŹWIAK A. OWCZAREK P, PROCHOWSKI L, ŚWIDERSKI A. **Analysis of the impact of the use time of n1 motor vehicles on the economic efficiency of their maintenance**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 121–129, http://dx.doi.org/10.17531/ein.2020.1.14.

The efficiency of operation of motor vehicles with a DMC (Permissible Laden Mass) <3.5 tonnes is considered. These are vehicles belonging motor vehicles of category N1, usually referred to as delivery vehicles. The results of observations on the implementation of transport orders in 7 transport companies from the MŚP (Small and Middle-size Companies) sector were used to conduct the effectiveness analysis. The research group covered 24 vehicles that implementation transport orders in the urban zone and in the immediate vicinity of the city. Information was collected on a monthly basis.During the analysis of economic efficiency the income measures (absolute and relative) were used. The calculations were carried out using the model of the vehicle operation process in the form of a neural network, in which a set of 12 input variables and 3 output variables were taken into account. Using the Statistica 13.3 computer program and defining the group and factors describing the process of implementation of individual transport tasks, the developed neural network model enabled searching for the impact of selected operational factors on the economic efficiency of N1 category cars.The calculations showed a significant impact of the number of vehicle days in a month, the weight of the load, as well as the time of year. The obtained calculation results showed the specific features of the impact of the number of working days on revenue in a transport company. The increase in the number of working days favors the increase in income in a limited way, and this restriction depends, among others since the time of year.

JÓŹWIAK A. OWCZAREK P, PROCHOWSKI L, ŚWIDERSKI A. **Badanie wpływu czasu wykorzystania samochodów kategorii n1 na efektywność ekonomiczną ich eksploatacji**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 121–129, http://dx.doi.org/10.17531/ein.2020.1.14.

Rozważa się efektywność eksploatacji samochodów ciężarowych o DMC < 3,5 tony. Są to pojazdy należące do kategorii N1 (według Dyrektywy 2007/46/WE) zwykle nazywane samochodami dostawczymi. Do prowadzonej analizy efektywności wykorzystano wyniki obserwacji z realizacji zleceń przewozowych w 7 firmach transportowych z sektora MŚP. Grupa badawcza objęła 24 pojazdy, które wykonywały zadania transportowe w strefie miejskiej i w najbliższym otoczeniu miasta. Informacje gromadzono w cyklach miesięcznych. Podczas analizy efektywności ekonomicznej zastosowano kilka miar przychodu (bezwzględny i względny). Obliczenia prowadzono przy wykorzystaniu modelu procesu eksploatacji pojazdów w postaci sieci neuronowej, w której brano pod uwagę zbiór 12 zmiennych wejściowych i 3 zmienne wyjściowe. Stosując program komputerowy Statistica 13.3 oraz zdefiniowanie grupy i czynniki opisujące proces realizacji poszczególnych zadań transportowych, opracowany model sieci neuronowej umożliwił poszukiwanie wpływu wybranych czynników eksploatacyjnych na efektywność ekonomiczną samochodów kategorii N1. Przeprowadzone obliczenia pokazały istotny wpływ liczby dni pracy pojazdów w miesiącu, masę ładunku, a także porę roku. Uzyskane wyniki obliczeń pokazały specyficzne cechy wpływu liczby dni pracy na przychód w firmie transportowej. Wzrost liczby dni pracy sprzyja wzrostowi przychodu w sposób ograniczony, a to ograniczenie zależy m.in. od pory roku.

PIELECHA I, PIELECHA J. **Simulation analysis of electric vehicles energy consumption in driving tests**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 130–137, http://dx.doi.org/10.17531/ein.2020.1.15.

The assessment of energy flow through electric vehicle systems makes estimating their energy consumption possible. The article presents analyzes of the energy consumption of electric vehicles in selected driving tests (NEDC, WLTC and in real traffic conditions – RDC test) in relation to the vehicles different curb weight. The use of electric motors was also analyzed, providing their operating ranges, data of the energy flow in batteries and the change in their charge level. Simulation tests and analyzes were carried out using the AVL Cruise software. It was found that despite similar vehicle energy consumption values in NEDC and RDC testing, there are significant differences in energy flow in vehicle subsystems. The changes in the battery charge level per 100 km of test drive are similar in both the WLTC and RDC tests (6% difference); for the NEDC test, this difference is the greatest at 25% (compared to the previous tests). The energy consumption of electric vehicles depends significantly on the test itself; the values obtained in the tests are in the ranges of 10.1–13.5 kWh/100 km (NEDC test); 13–15 kWh/100 km (WLTC test) and 12.5–16.2 kWh/100 km in the RDC test. The energy consumption values in the NEDC and WLTC tests, compared to the RDC test, are approximately 20% and 10% lower, respectively. Increasing the vehicle mass increases the energy consumption (increasing the vehicle mass by 100 kg was found to increase the energy consumption by 0.34 kWh/100 km).

PIELECHA I, PIELECHA J. **Symulacyjna analiza energochłonności pojazdów elektrycznych w testach badawczych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 130–137, http://dx.doi.org/10.17531/ein.2020.1.15.

Ocena przepływu energii przez układy pojazdów elektrycznych umożliwia oszacowanie ich energochłonności. W artykule przedstawiono analizy dotyczące zużycia energii pojazdów elektrycznych w wybranych testach jezdnych (NEDC, WLTC oraz w rzeczywistych warunkach ruchu – test RDC) w odniesieniu do zróżnicowanej masy pojazdów. Analizie poddano również wykorzystanie silników elektrycznych, przedstawiając mapy ich pracy, wielkości przepływu energii w akumulatorach oraz stopień zmiany ich naładowania. Badania i analizy symulacyjne wykonano z wykorzystaniem oprogramowania AVL Cruise. Stwierdzono, że mimo podobnych wartości energochłonności pojazdów w testach badawczych NEDC oraz RDC, to występują znaczące różnice przepływu energii w układach akumulacji pojazdów. Zmiany stopnia naładowania akumulatora odniesione do 100 km testu są zbliżone w testach WLTC oraz RDC (różnica 6%); dla testu NEDC różnica ta wynosi maksymalnie 25% (w odniesieniu do poprzednich testów). Energochłonność pojazdów elektrycznych jest silnie zależne od testu badawczego; wartości uzyskane w testach kształtują się na poziomie 10,1–13,5 kWh/100 km (test NEDC); 13–15 kWh/100 km (test WLTC) oraz 12,5–16,2 kWh/100 km w teście RDC. Wartości energochłonności w testach NEDC oraz WLTC są odpowiednio mniejsze o około 20% i 10% w odniesieniu do testu RDC. Zwiększenie masy pojazdu zwiększa zużycie energii (zwiększenie o 100 kg masy pojazdu zwiększa zużycie energii o 0,34 kWh/100 km).

KŁOSOWSKI G, RYMARCZYK T, KANIA K, ŚWIĆ A, CIEPLAK T. **Maintenance of industrial reactors supported by deep learning driven ultrasound tomography**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 138–147, http://dx.doi.org/10.17531/ein.2020.1.16.

Monitoring of industrial processes is an important element ensuring the proper maintenance of equipment and high level of processes reliability. The presented research concerns the application of the deep learning method in the field of ultrasound tomography (UST). A novel algorithm that uses simultaneously multiple classification convolutional neural networks (CNNs) to generate monochrome 2D images was developed. In order to meet a compromise between the number of the networks and the number of all possible outcomes of a single network, it was proposed to divide the output image into 4-pixel clusters. Therefore, the number of required CNNs has been reduced fourfold and there are 16 distinct outcomes from single network. The new algorithm was first verified using simulation data and then tested on real data. The accuracy of image reconstruction exceeded 95%. The results obtained by using the new CNN clustered algorithm were compared with five popular machine learning algorithms: shallow Artificial Neural Network, Linear Support Vector Machine, Classification Tree, Medium k-Nearest Neighbor classification and Naive Bayes. Based on this comparison, it was found that the newly developed method of multiple convolutional neural networks (MCNN) generates the highest quality images.

KŁOSOWSKI G, RYMARCZYK T, KANIA K, ŚWIĆ A, CIEPLAK T. **Eksploatacja reaktorów przemysłowych ze wspomaganiem tomografii ultradźwiękowej i algorytmów głębokiego uczenia**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 138–147, http://dx.doi.org/10.17531/ein.2020.1.16.

Monitorowanie procesów przemysłowych jest ważnym elementem zapewniającym właściwą eksploatację urządzeń i wysoki poziom niezawodności procesów. Prezentowane badania dotyczą zastosowania metod głębokiego uczenia w obszarze eksploatacji zbiornikowych reaktorów przemysłowych. W procesach przemysłowych opartych na reakcjach chemicznych zachodzących wewnątrz procesowej tomografii ultradźwiękowej (UST). Opracowano nowatorski algorytm wykorzystujący jednocześnie wiele klasyfikacyjnych splotowych sieci neuronowych (CNN) do generowania monochromatycznych obrazów 2D. Aby osiągnąć kompromis między liczbą sieci a liczbą wszystkich możliwych wyników pojedynczej sieci, zaproponowano podział obrazu wyjściowego na klastry 4-pikselowe. W związku z tym liczba wymaganych CNN została czterokrotnie zmniejszona i istnieje 16 różnych wyników z jednej sieci. Nowy algorytm został najpierw zweryfikowany przy użyciu danych symulacyjnych, a następnie przetestowany na danych rzeczywistych. Dokładność rekonstrukcji obrazu przekroczyła 95%. Wyniki uzyskane przy użyciu nowego algorytmu klastrowego CNN zostały porównane z pięcioma popularnymi algorytmami uczenia maszynowego: płytką sztuczną siecią neuronową, maszyną liniowego wektora wsparcia, drzewem klasyfikacji, klasyfikacją średniego k-najbliższego sąsiada i naiwnym Bayesem. Na podstawie tego porównania stwierdzono, że nowo opracowana metoda wielu splotowych sieci neuronowych (MCNN) generuje obrazy o najwyższej jakości.

MIELCZAREK M, SŁOWIK M, ANDRZEJCZAK K. **The assessment of influence of styrene-butadiene-styrene elastomer's content on the functional properties of asphalt binders**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 148–153, http://dx.doi.org/10.17531/ein.2020.1.17.

This paper discusses the issue of improving the functional properties of road asphalt pavements by modifying bituminous binder with SBS copolymer. The main purpose of the paper is to assess the resistance to permanent deformations and the temperature susceptibility of polymer-modified road asphalt binders, which are most commonly used in the upper layers of road and airport pavements. The bitumens subject to the study originate from various crude oil deposits (Russian and Venezuelan). They were modified in laboratory conditions with a concentrated additive with the known content of the SBS copolymer of 9%. The result was a asphalt binder containing the known percentage of the SBS copolymer of 1.5%, 3.0%, 4.5% and 6%. The rheological properties of the tested bitumens were determined by use of a dynamic shear rheometer (DSR), and with the application of the sinusoidal variable load, in the broad test temperature spectrum (from 40°C to 100°C). The analysis of the values of the dynamic shear modulus |G*| of all the studied bitumens shows that the increase in the content of SBS copolymer in the tested binder increases the value of |G*|, which may result in higher resistance to permanent deformations of road pavements caused by repeated traffic loads, especially in the case of pavements operated at high temperatures. The asphalt mixtures resistance to rutting is one of the basic parameters related to road pavement service-life, affecting both the safety and driving comfort of users.

MIELCZAREK M, SŁOWIK M, ANDRZEJCZAK K. **Ocena wpływu zawartości elastomeru styren-butadien-styren na właściwości funkcjonalne lepiszczy asfaltowych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 148–153, http://dx.doi.org/10.17531/ein.2020.1.17.

Tematyka pracy związana jest z zagadnieniem polepszenia właściwości funkcjonalnych drogowych nawierzchni asfaltowych poprzez modyfikację lepiszcza asfaltowego kopolimerem SBS. Głównym celem pracy jest ocena odporności na odkształcenia trwałe oraz wrażliwości na zmiany temperatury asfaltów drogowych modyfikowanych polimerami, które są najczęściej używane w wierzchnich warstwach konstrukcji nawierzchni drogowych i lotniskowych. Przedmiotem badań były asfalty pochodzące z różnych złóż ropy naftowej (rosyjskiej i wenezuelskiej). Asfalty te poddano modyfikacji w warunkach laboratoryjnych z dodatkiem koncentratu o znanej zawartości kopolimeru SBS równej 9%. Otrzymano w ten sposób lepiszcza asfaltowe o znanej zawartości kopolimeru SBS równej 1,5%; 3,0%; 4,5% oraz 6%. Właściwości reologiczne badanych asfaltów oznaczono z użyciem reometru dynamicznego ścinania DSR stosując w testach obciążenie sinusoidalnie zmienne, w szerokim zakresie temperatury pomiarowej (od 40°C do 100°C). Analizując wartości dynamicznego modułu ścinania |G*| wszystkich badanych asfaltów można stwierdzić, iż wzrost zawartości kopolimeru SBS w badanym lepiszczu zwiększa wartość |G*|, co może skutkować większą odpornością na odkształcenia trwałe nawierzchni drogowej spowodowane wielokrotnie powtarzającymi się obciążeniami ruchem pojazdów, w szczególności w przypadku nawierzchni eksploatowanej w wysokiej temperaturze Odporność mieszanek mineralno-asfaltowych (MMA) na powstawanie kolein jest jednym z podstawowych parametrów związanych z eksploatacją nawierzchni drogowych, wpływając zarówno na bezpieczeństwo, jak i komfort jazdy użytkowników.

SZUMSKA E, JURECKI R, PAWEŁCZYK M. **Evaluation of the use of hybrid electric powertrain system in urban traffic conditions**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 154–160, http://dx.doi.org/10.17531/ein.2020.1.18.

The conditions of use of the vehicle significantly affect the performance results. Traffic conditions in a specific city directly affect the consumption of energy, fuel and emissions of harmful compounds in exhaust fumes. Conduction of the measurements of a vehicle's performance parameters in operating conditions is very troublesome and is often not possible to realize. An alternative is to use the simulation programs. Vehicle simulation programs offer options related to vehicle models or drive unit components and allow development of new models. Based on the results of simulation testing, it is possible to analyse the level of fuel and energy consumption as well as emissions of harmful compounds in exhaust gases and the operating effectiveness of the drive system in the speed profile. The paper presents the evaluation of the effectiveness of using hybrid electric drive system in passenger cars in medium-sized city traffic conditions using the Kielce example. The simulation tests were based on the speed profiles recorded during real-world test drives in various times of the day. The simulation results were used to conduct an analysis of fuel consumption and pollutant emissions recorded by conventional and hybrid vehicles.

SZUMSKA E, JURECKI R, PAWEŁCZYK M. **Ocena zastosowania napędów hybrydowych w warunkach ruchu miejskiego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 154–160, http://dx.doi.org/10.17531/ein.2020.1.18.

Warunki użytkowania pojazdu mają znaczący wpływ na parametry eksploatacyjne pojazdu. Warunki ruchu w określonym mieście bezpośrednio wpływają na zużycie energii, paliwa i poziom emisji szkodliwych związków zawartych w spalinach. Przeprowadzenie pomiarów parametrów eksploatacyjnych pojazdu w warunkach rzeczywistych jest kłopotliwe i często niemożliwe do zrealizowania. Alternatywą jest wykorzystanie symulacji komputerowych. Programy do symulacji pojazdów oferują, między innymi, modele pojazdów lub komponentów układu napędowego oraz pozwalają na opracowanie nowych modeli. Na podstawie wyników badań symulacyjnych możliwa jest analiza poziomu zużycia paliwa, energii, emisji szkodliwych związków zawartych w spalinach oraz efektywności pracy układu napędowego w profilu prędkości. W niniejszej pracy przedstawiono ocenę efektywności zastosowania napędów hybrydowych w samochodach osobowych w warunkach ruchu miasta średniej wielkości na przykładzie Kielc. Do badań symulacyjnych wykorzystano profile prędkości, zarejestrowane podczas rzeczywistych przejazdów w różnych porach dnia. Na podstawie wyników symulacji przeprowadzono analizę zużycia paliwa oraz emisji zanieczyszczeń, zarejestrowanych dla pojazd z napędem konwencjonalnym oraz pojazdów z napędem hybrydowym.

ZHANG X, ZHAO J. **Compound fault detection in gearbox based on time synchronous resample and adaptive variational mode decomposition**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 161–169, http://dx.doi.org/10.17531/ein.2020.1.19.

Compound fault detection of gearboxes is an ambitious matter considering its interconnection and complication. An innovative means for compound fault detection based on time synchronous resample (TSR) and adaptive variational mode decomposition (AVMD) is put forward in this work. TSR used in the method can enhance fault signals of synchronous shaft gears by eliminating signal components independent of synchronous shaft. Therefore, the TSR is used to separate the synchronous shaft signal corresponding to the gear fault from the raw compound fault signal. Then a series of mode components are obtained by decomposing the synchronous shaft signals of all faults by AVMD. The variational mode decomposition (VMD) can overcome the mode aliasing problem of empirical mode decomposition (EMD), but the decomposition effect of VMD is affected by its parameter setting. Thus, the paper proposes an AVMD algorithm based on whale optimization algorithm (WOA). In the AVMD, the WOA is used to optimizes the parameters of the VMD. After AVMD decomposition, the correlated kurtosis of the mode components obtained by AVMD decomposition is calculated. Then the mode components with the maximum correlated kurtosis are selected to carry out envelope analysis. Finally, the compound fault feature can be found from the envelope spectrum to get the diagnosis results. In order to test the validity of the proposed method, a compound fault experiment is implemented in a gearbox. Through the analysis of the experimental data, it is proved that the method shows a good performance in the compound fault detection of gearbox.

ZHANG X, ZHAO J. **Wykrywanie złożonych błędów przekładni na podstawie synchronicznego próbkowania wtórnego oraz adaptacyjnej metody wariacyjnej dekompozycji modalnej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 161–169, http://dx.doi.org/10.17531/ein.2020.1.19.

Wykrywanie złożonych błędów przekładni stanowi trudne zagadnienie ze względu na ich skomplikowany charakter i powiązania wewnętrzne. W pracy zaproponowano nowatorską metodę wykrywania błędów złożonych opartą na synchronicznym próbkowaniu wtórnym (TSR) oraz adaptacyjnej metodzie wariacyjnej dekompozycji modalnej (AVMD). TSR pozwala wzmacniać sygnały błędów występujących w synchronicznych przekładniach walcowych, dzięki eliminacji składowych sygnału niezwiązanych z działaniem wału synchronicznego. Dlatego też w przedstawionych badaniach, TSR wykorzystano do wyodrębnienia sygnału wału synchronicznego odpowiadającego błędowi przekładni, z surowego sygnału błędu złożonego. Następnie wszystkie sygnały błędu wału synchronicznego poddano dekompozycji za pomocą AVMD, dzięki czemu otrzymano szereg składowych modalnych. Wariacyjna dekompozycja modalna (VMD) pozwala uniknąć problemu aliasingu, który występuje w przypadku empirycznej dekompozycji modalnej (EMD), przy czym efekt dekompozycji zależy od ustawień parametrów. Dlatego w artykule zaproponowano adaptacyjny algorytm VMD oparty na algorytmie optymalizacji wielorybów (WOA), który optymalizuje parametry VMD. Następnym krokiem po dekompozycji AVMD, było obliczenie skorelowanej kurtozy składowych modalnych otrzymanych na drodze tej dekompozycji. Składniki modalne o najwyższych wartościach skorelowanej kurtozy wykorzystano do przeprowadzenia analizy obwiedni. Błąd złożony wykrywano na podstawie widma obwiedni. Skuteczność proponowanej metody sprawdzono przeprowadzając doświadczenie na przekładni, w której występował błąd złożony. Wyniki eksperymentu pokazują, że proponowane podejście stanowi skuteczną metodę wykrywania złożonych błędów.

LI J, WANG Z, REN Y, YANG D, LV X. **A novel reliability estimation method of multi-state system based on structure learning algorithm**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 170–178, http://dx.doi.org/10.17531/ein.2020.1.20.

Traditional reliability models, such as fault tree analysis (FTA) and reliability block diagram (RBD), are typically constructed with reference to the function principle graph that is produced by system engineers, which requires substantial time and effort. In addition, the quality and correctness of the models depend on the ability and experience of the engineers and the models are difficult to verify. With the development of data acquisition, data mining and system modeling techniques, the operational data of a complex system considering multi-state, dependent behavior can be obtained and analyzed automatically. In this paper, we present a method that is based on the K2 algorithm for establishing a Bayesian network (BN) for estimating the reliability of a multi-state system with dependent behavior. Facilitated by BN tools, the reliability modeling and the reliability estimation can be conducted automatically. An illustrative example is used to demonstrate the performance of the method.

LI J, WANG Z, REN Y, YANG D, LV X. **Nowatorska metoda oceny niezawodności systemów wielostanowych w oparciu o algorytm uczenia struktury**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 170–178, http://dx.doi.org/10.17531/ein.2020.1.20.

Tradycyjne modele niezawodności, takie jak analiza drzewa błędów (FTA) czy schemat blokowy niezawodności (RBD), buduje się zazwyczaj w oparciu o tworzone przez inżynierów systemowych schematy zasad działania systemu, których przygotowanie wymaga dużych nakładów czasu i pracy. Jakość i poprawność tych modeli zależy od umiejętności i doświadczenia inżynierów, a same modele są trudne do zweryfikowania. Dzięki rozwojowi technik akwizycji i eksploracji danych oraz modelowania systemów, dane operacyjne złożonego systemu uwzględniające jego zależne, wielostanowe zachowania mogą być pozyskiwane i analizowane automatycznie. W artykule przedstawiono metodę konstrukcji sieci bayesowskiej (BN) opartą na algorytmie K2, która pozwala na ocenę niezawodności systemu wielostanowego o zachowaniach zależnych. Dzięki narzędziom BN, modelowanie i szacowanie niezawodności może odbywać się automatycznie. Działanie omawianej metody zilustrowano na podstawie przykładu.

Article citation info:
KILIKEVIČIUS A, RIMŠA V, RUCKI M. Investigation of influence of aircraft propeller modal parameters on small airplane performance. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 1–5, http://dx.doi.org/10.17531/ein.2020.1.1.

Arturas KILIKEVIČIUS
Vytautas RIMŠA
Mirosław RUCKI

INVESTIGATION OF INFLUENCE OF AIRCRAFT PROPELLER MODAL PARAMETERS ON SMALL AIRPLANE PERFORMANCE

BADANIA WPŁYWU MODALNYCH PARAMETRÓW ŚMIGŁA NA ZACHOWANIE MAŁEGO SAMOLOTU

The aim of the current paper is to investigate a small airplane model propeller of class F2D according to requirements of Fédération Aéronautique Internationale (FAI, or World Air Sports Federation). In some cases, practical tests show that F2D models with flexible propellers produce specific extra noise and increase flight speed in comparison with "rigid" propellers. Therefore, the following hypothesis could be proposed: flexible characteristics of the increased noise are related to the resonant eigenfrequencies of the propeller. The operating range of the F2D class propeller (28,000-35,000 rpm) is close to or equal to the eigenfrequency resonance. The current investigation addresses dynamic/flexible vibrations of elastic propeller during engine run and researches dynamic parameters of the propeller as well as the contribution of these parameters to the model flight characteristics. To resolve this type of a problem, a stand, which allows completing a physical investigation of flexible propeller vibration modes and dynamic characteristics was created.

***Keywords**: aircraft propeller, dynamics, vibration, operational modal analysis.*

Celem artykułu jest przedstawienie wyników badań śmigła małego modelu samolotu zaliczanego do klasy F2D (według klasyfikacji Fédération Aéronautique Internationale, FAI). W niektórych przypadkach testy wykazały, że modele F2 z giętkimi śmigłami, w porównaniu do śmigieł sztywnych, wydają dodatkowy hałas i zwiększają prędkość samolotu. Dlatego wysunięto hipotezę, że elastyczne charakterystyki zwiększonego hałasu są powiązane z rezonansem częstotliwości własnych śmigła. Zakres pracy śmigła klasy F2D (28 000-35 000 obr/min) jest zbliżony do jego częstotliwości własnych. Badania dotyczą elastycznych wibracji dynamicznych śmigła giętkiego w czasie rozruchu silnika i są nakierowane na wyznaczanie parametrów dynamicznych i ich wpływu na charakterystyki lotu modelu. Wykonano i opisano stanowisko, na którym przeprowadzono testy modalne drgań giętkiego śmigła. Na tej podstawie uzyskano charakterystyki dynamiczne.

***Słowa kluczowe**: śmigło lotnicze, dynamika, wibracje, analiza modalna.*

1. Introduction

The phenomenon of aero-elasticity is a significant field of research in aviation. The studies show the interactions among inertial, elastic and aerodynamic forces that occur when flexible body is exposed by the fluid flow [1, 7, 25]. Aero elasticity draws on the study of fluid mechanics, solid mechanics, structural dynamics and dynamical systems. The conventional classification of aero-elasticity includes two fields: static aero elasticity, which deals with the static or steady response of an elastic body to a fluid flow; and dynamic aero elasticity, which deals with the body's dynamic (typically vibrational) response.

The phenomenon of dynamic aero elastic stability is formed by vortex oscillations and has a significant influence on the propeller working process. The turbulence is caused by the rotating parts of the turbine (propeller or gas turbine or engine rotors). Whirl flutter instability is a specific type of aero elastic layer instability that can occur in turbo propellers [8, 15]. The propeller is intended to change motion of the engine's shaft into the traction power of the aircraft. It conversion takes place due to differential pressure which is created in the front and in the aft of the propeller blades. Therefore, it leads to high vortex oscillation effect especially near the tip of the propeller. That may be an issue especially in case of small aircrafts [20].

The vortex oscillation effect is the interaction of solid body motion and aerodynamic forces (FSI), which are very common for such elements as the gas turbine engine rotor or propeller. Rotating mass increases the number of degrees of freedom and creates extra forces and moments [5, 24]. In some cases, vortex oscillations result in highly large and unstable aerodynamic forces and moments thus affecting the propeller or the aircraft as such and can lead to the distortion of the construction. Moreover, the vibrations in the airplane usually have a negative effect, so many researches are aimed to reduction their impact on the construction [11-14, 22]. This paper is focused on the resonance phenomena in F2D model class, where resonance may affect durability performance of the construction. On the other hand,

it helps to increase propeller trust, which is the key factor [16, 23]. Investigation of propeller blade deflection and frequencies analysis demonstrated, that higher model speed was achievable.

However, in some cases experimental tests show that F2D models with flexible propellers have specific extra noise and increase the speed of flight in comparison with "rigid" propellers. The test demonstrated up to 2.2% higher speed, which decreased 1 km distance flight time from 23.0 s down to 22.5 s. This is a remarkable improvement of the F2D model flight speed within restrictions set by FAI (Fédération Aéronautique Internationale, or World Air Sports Federation) regulations.

This article discusses the specifications of F2D class aircraft propellers and their dynamic characteristics. The article introduces an experimental propeller test stand that allows to determine the dynamic characteristics of a flexible propeller. Experimental studies have shown that when using flexible propellers with first and second resonant frequencies within the working range of the aircraft, better flight parameters are achieved.

2. Modal Analysis issue

Operational modal analysis (OMA) is a sort of 'inverse' problem where one is interested in gaining knowledge about the instrumented structure based on measured response [26]. The method is economical and feasible, so it is becoming a common practice in full-scale vibration testing worldwide [2]. There are two main types of OMA algorithms, one operating in time domain and other in frequency domain [6]. In the first group can be found approaches based on the dynamic state equations and stochastic subspace identification (SSI), which provides fairly accurate estimates of the low frequency modes under normal operating conditions but may be further improved [3, 4]. The second group is represented with various methods of the system responses decomposition in the frequency domain, such as Basic Frequency Domain (BFD), Frequency Domain Decomposition (FFD), or Enhanced Frequency Domain Decomposition (EFFD) [9, 17, 19]. The key requirements that the system under analysis must meet are the following: the structure is unchanging in time; low inhibition; well-separated specific frequencies; structure-induced excitation is stationary broadband noise.

Grosel and co-authors [10] provided the following relation between excitation $x(t)$ and response $y(t)$:

$$\left[G_{yy}(j\omega)\right]=\left[H(j\omega)\right]^{*}\left[G_{yy}(j\omega)\right]\left[H(j\omega)\right]^{T} \tag{1}$$

where the brackets [] represented a matrix, $[G]$ - the power spectral density matrix with index xx representing the input and the index yy denoting the output, $[H]$ - represents the matrix of the frequency response function, T – transposition, and the asterisk * means the complex conjugate.

They further proposed the splitting of the response matrix using the Singular Value Decomposition as follows:

$$\left[G_{yy}(j\omega_i)\right]=[\Phi_i][S_i][\Phi_i]^{H} \tag{2}$$

where S_i – diagonal of the Singular Value Decomposition, Φ_i – unitary matrix containing vectors that are proportional to their own value vectors. The diagonal elements of the S_i matrix contain information about their frequencies.

Dynamic behavior of the structure in a given frequency range can be modeled as a set of individual modular vibrations. It is assumed that the structure acts as a linear, time-changing system. Each resonant frequency can be described by the following parameters: self-frequency or resonant frequency, (modal) shape, and damping factor. These parameters are called modal parameters. Using modal parameters for structural modeling, vibration problems caused by these resonances (mods) can be assessed and analyzed. In addition, the model can be used to provide possible solutions to individual issues. Modal parameters can be derived from Frequency Response Functions (FRF) measurements consisting of one or more reference positions and multiple measurement positions required to describe the model behavior.

3. Experimental equipment for assessment of dynamic parameters

In the current investigation, the examined objects were the air model propellers. For the test three different propellers were used: (1) F2C class team race propeller made from carbon rigid sample, (2) F2D class propeller called "C1", made from fiber glass, one of the most popular propellers (length 155mm; pitch 80) which is used in various weather conditions, and (3) F2D class propeller called "ZM", made from fiber glass, which is known from the practical perspective as a resonant propeller (length 155 mm, pitch 80 mm). The propellers are shown in Fig. 1.

The "Brüel & Kjær" and "Lion Precision" measuring instruments were used to measure vibration parameters. In Fig. 2, the main subunits of the devices are shown: 1) Excitation Vibrator 4810 with

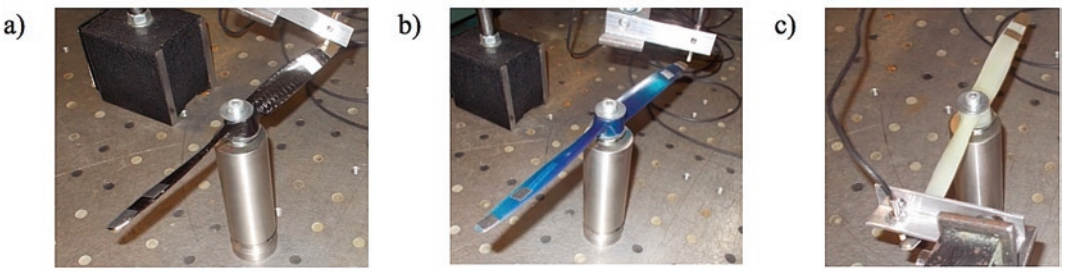


Fig. 1. The propellers under tests: a) F2C class carbon sample; b) F2D class "C1" fiber glass; c) F2D class "ZM" fiber glass

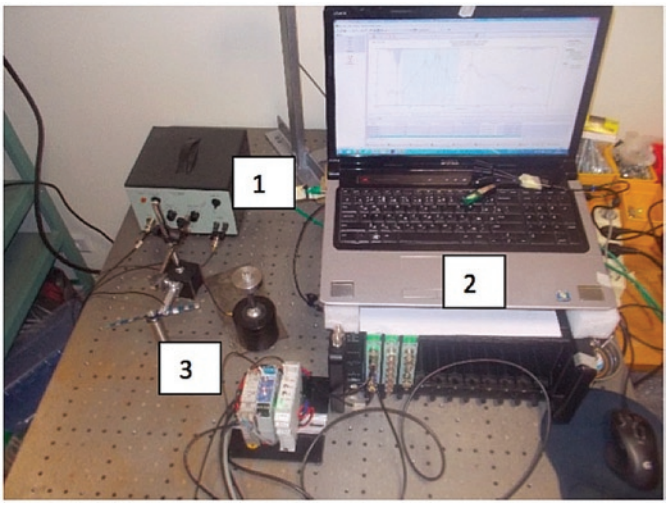


Fig. 2. Apparatus for the vibration tests of propellers

amplifier, 2) Portable measurement processing, storage and control equipment No. 3660D with PC, and 3) Displacement sensors "Lion Precision" U3B and U20B with amplifier. Data flow is presented in the chart in Fig 3.

The measurements of the propellers' vibration were performed when the propeller was rigidly fixed to the rod in the same way as it is shown in the model. The rod was fixed to the vibration frequency generator. Vibration test was done by the use "sweep sine" method in the range of frequency from 100 up to 2000 Hz.

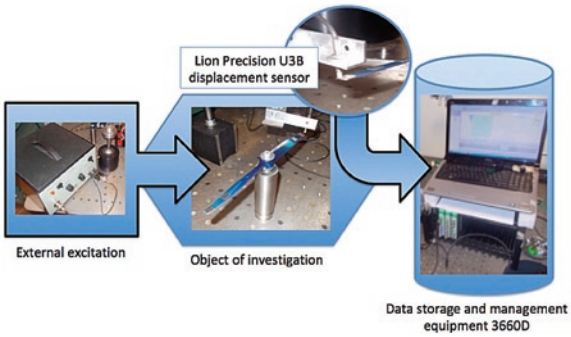


Fig. 3. Measurement data flow chart

4. Results and discussion

Fig. 4 shows the graphs of power spectra of the three investigated propellers obtained using Frequency Domain Decomposition (FDD) technique, which allow for estimating natural frequencies and mode shapes of the structural system [18]. The areas of modal domain are marked with green colors with estimated modes shown in red squares. The natural frequencies of investigated propellers are presented in Table 1. In the Table 2, the shapes obtained from Operational Modal Analysis (OMA) and Blade Element Momentum (BEM) are shown. Classical momentum theory can be applied to the ideal propeller, but blade element theory is used for more detailed analysis of its exploitation characteristics [21].

The presented results demonstrated that F2C class carbon propeller had its resonating frequencies close to each other in the span of ca. 150 Hz between 905 and 1050 Hz. The largest, almost six times wider span between the modes was found in the "ZM" propeller, which was ca. 850 Hz between 400 and 1250 Hz. However, only in case of F2D class "C1" propeller made out of fiber glass, its resonant frequency was covered by the operating conditions. Namely, F2D class aircraft models had operating range 28,000-35,000 rpm, corresponding with 467-583.3 Hz. Most probably, its improved exploitation characteristics can be attributed to that fact.

Both OMA and BEM models indicate larger deformations of the propeller blade on its tips. This phenomenon can provide explana-

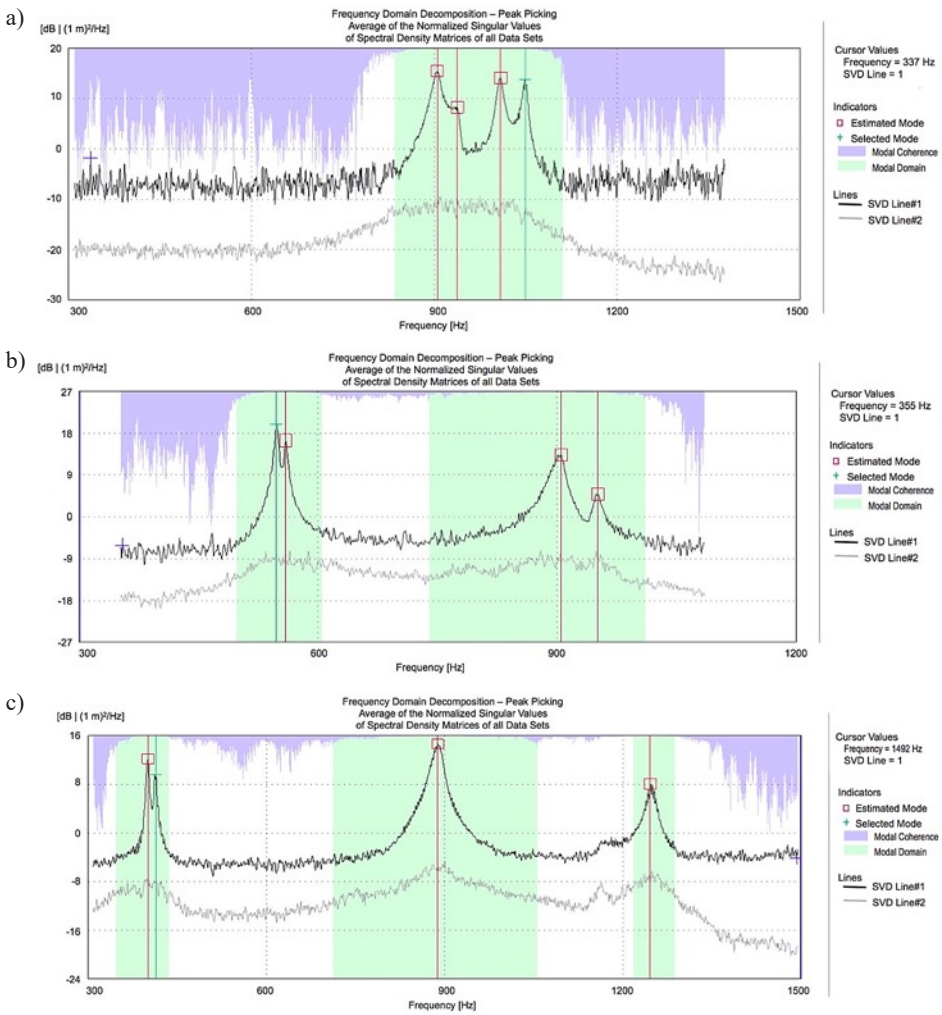
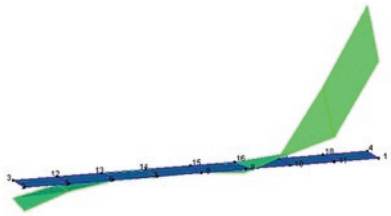
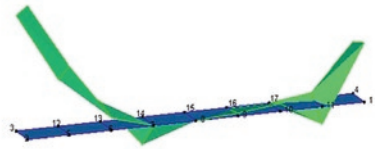
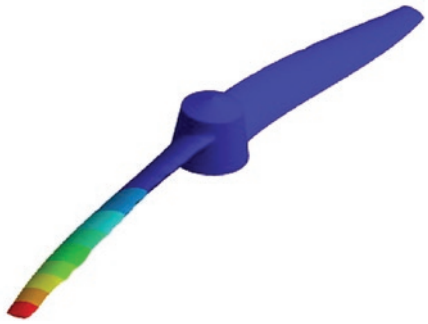


Fig. 4. The power spectral density functions of the investigated propellers obtained from FDD technique: a) F2C class carbon sample; b) F2D class "C1" fiber glass; c) F2D class "ZM" fiber glass

Table 1. The natural frequencies of investigated propellers

Mode	Frequency [Hz]		
	F2C class carbon propeller	F2D class glass propeller (C1)	F2D class glass propeller (ZM)
Mode 1	905	548	402
Mode 2	938	560	413
Mode 3	1009	905	889
Mode 4	1049	951	1253

Table 2. Deformations of the blades in modes 1 and 2 obtained from OMA and BEM models

	OMA shapes	BEM shapes
Mode 1		
Mode 2		

tion why the differences between two analyzed modes 1 and 2 are so small, namely, between 2.2% and 3.6% for any type of the tested propellers.

5. Conclusion

The current analysis of the dynamic characteristics of propellers used in F2D class aircraft models engaged both experimental measurements and operational modal analysis. Their operating range was between 28,000 and 35,000 rpm. Three different propellers of different shape and structural materials were investigated in terms of modal frequency parameters. The investigation revealed the frequencies of the propellers, which were: 1.: 905, 938, 1009 and 1049 Hz for F2C class carbon racing, 548, 560, 905 and 951 Hz for F2D class (C1) fiber glass propeller, and 402, 413, 889 and 1253 Hz for F2D class (ZM) fiber glass propeller, respectively.

The first two resonant frequencies (548 and 560 Hz) of the C1 propeller have been found in the frequency range corresponding with the F2D aircraft models operating conditions, i.e. 467-583.3 Hz corresponding to the propeller rotational speed 28,000-35,000 rpm. To that fact may be attributed its higher speed.

Acknowledgement

The research work had no specific financial support. International cooperation was possible thank to the Erasmus+ program.

References

1. Aref P, Ghoreyshi M, Jirasek A, Satchell M, Bergeron K. Computational Study of Propeller-Wing Aerodynamic Interaction. Aerospace 2018; 5(3): 79, https://doi.org/10.3390/aerospace5030079.
2. Au S-K, Brownjohn J, Mottershead J E. Quantifying and managing uncertainty in operational modal analysis. Mechanical Systems and Signal Processing 2018; 102: 139-157, https://doi.org/10.1016/j.ymssp.2017.09.017.
3. Bajrić A, Høgsberg J, Rüdinger F. Evaluation of damping estimates by automated Operational Modal Analysis for offshore wind turbine tower vibrations. Renewable Energy 2018; 116: 153-163, https://doi.org/10.1016/j.renene.2017.03.043.
4. Philip J G, Jain T. An improved Stochastic Subspace Identification based estimation of low frequency modes in power system using synchrophasors. International Journal of Electrical Power & Energy Systems 2019; 109: 495-503, https://doi.org/10.1016/j.ijepes.2019.01.030.
5. Başak H, Prempain E. Switched fault tolerant control for a quadrotor UAV. IFAC-Papers On Line 2017; 50(1): 10363-10368, https://doi.org/10.1016/j.ifacol.2017.08.1686.

6. Brandt A. A signal processing framework for operational modal analysis in time and frequency domain. Mechanical Systems and Signal Processing 2019; 115: 380-393, https://doi.org/10.1016/j.ymssp.2018.06.009.
7. Bronstein M, Feldman E, Vescovini R, Bisagni C. Assessment of dynamic effects on aircraft design loads: The landing impact case. Progress in Aerospace Sciences 2015; 78: 131-139, https://doi.org/10.1016/j.paerosci.2015.06.003.
8. Čečrdle J. Whirl Flutter of Turboprop Aircraft Structures. Amsterdam: Elsevier, 2015.
9. Goyal D, Pabla B S. The Vibration Monitoring Methods and Signal Processing Techniques for Structural Health Monitoring: A Review. Archives of Computational Methods in Engineering 2016; 23(4): 585-594, https://doi.org/10.1007/s11831-015-9145-0.
10. Grosel J, Sawicki W, Pakos W. Application of Classical and Operational Modal Analysis for Examination of Engineering Structures. Procedia Engineering 2014; 91: 136-141, https://doi.org/10.1016/j.proeng.2014.12.035.
11. Jurevicius M, Skeivalas J, Kilikevicius A, Turla V. Vibrational analysis of length comparator. Measurement 2017; 103: 10-17, https://doi.org/10.1016/j.measurement.2017.02.010.
12. Kilikevičius A, Čereška A, Kilikevičienė K. Analysis of external dynamic loads influence to photovoltaic module structural performance. Engineering Failure Analysis 2016; 66: 445-454, https://doi.org/10.1016/j.engfailanal.2016.04.031.
13. Kilikevicius A, Jurevicius M, Skeivalas J, Kilikeviciene K, Turla V. Vibrational analysis of angle measurement comparator. Signal, Image and Video Processing 2016; 10(7): 1287-1294, https://doi.org/10.1007/s11760-016-0956-8.
14. Kilikevičius A, Kasparaitis A. Dynamic research of multi-body mechanical systems of angle measurement. International Journal of Precision Engineering and Manufacturing 2017; 18(8): 1065-1073, https://doi.org/10.1007/s12541-017-0125-1.
15. Kim T, Lim J, Shin S, Kim D-H. Structural design optimization of a tiltrotor aircraft composite wing to enhance whirl flutter stability. Composite Structures 2013; 95: 283-294, https://doi.org/10.1016/j.compstruct.2012.08.019.
16. Kopecki T, Mazurek P, Lis T. The effect of the type of elements used to stiffen thin-walled skins of load-bearing aircraft structures on their operating properties. Experimental tests and numerical analysis. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (2): 164-170, https://doi.org/10.17531/ein.2016.2.2.
17. Nita G M, Mahgoub M A, Sharyatpanahi S G, Cretu N C, El-Fouly T M. Higher order statistical frequency domain decomposition for operational modal analysis. Mechanical Systems and Signal Processing 2017; 84, Part A: 100-112, https://doi.org/10.1016/j.ymssp.2016.07.004.
18. Pioldi F, Ferrari R, Rizzi E. Output-only modal dynamic identification of frames by a refined FDD algorithm at seismic input and high damping. Mechanical Systems and Signal Processing 2016; 68-69: 265-291, https://doi.org/10.1016/j.ymssp.2015.07.004.
19. Rizo-Patron S, Sirohi J. Operational Modal Analysis of a Helicopter Rotor Blade Using Digital Image Correlation. Experimental Mechanics 2017; 57(3): 367-375, https://doi.org/10.1007/s11340-016-0230-6.
20. Samolej S, Orkisz M, Rogalski T. The Airspeed Automatic Control Algorithm for Small Aircraft. In: Nawrat A, Bereska D, Jedrasiak K (Eds.). Advanced Technologies in Practical Applications for National Security. Cham: Springer, 2018: 157-168, https://doi.org/10.1007/978-3-319-64674-9_10.
21. Sforza P M. Theory of Aerospace Propulsion (Second Edition). Amsterdam: Elsevier, 2017, https://doi.org/10.1016/B978-0-12-809326-9.00013-0.
22. Šiaudinytė L, Kilikevičius A, Sabaitis D, Grattan K T V. Modal analysis and experimental research into improved centering-leveling devices. Measurement 2016; 88: 9-17, https://doi.org/10.1016/j.measurement.2016.01.044.
23. Stępień S, Szajnar S, Jasztal M. Problems of military aircraft crew's safety in condition of enemy counteraction. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (3): 441-446, https://doi.org/10.17531/ein.2017.3.15.
24. Tang Y-R, Xiao X, Li Y. Nonlinear dynamic modeling and hybrid control design with dynamic compensator for a small-scale UAV quadrotor. Measurement 2017; 109: 51-64, https://doi.org/10.1016/j.measurement.2017.05.036.
25. Teixeira P, Cesnik C E. Propeller Effects on the Dynamic Response of HALE Aircraft. 2018 AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Kissimmee, Florida, 2018, https://doi.org/10.2514/6.2018-1202.
26. Zhu Y-Ch, Au S-K, Brownjohn J. Bayesian operational modal analysis with buried modes. Mechanical Systems and Signal Processing 2019; 121: 246-263, https://doi.org/10.1016/j.ymssp.2018.11.022.

Arturas KILIKEVIČIUS
Institute of Mechanical Science,
Vilniaus Gedimino Technical University,
J. Basanaviciaus g. 28, Vilnius LT-03224, Lithuania

Vytautas RIMŠA
Department of Aviation Technologies,
Vilniaus Gedimino Technical University,
Linkmenų g. 28, Vilnius, LT-08217, Lithuania

Mirosław RUCKI
Faculty of Mechanical Engineering,
Kazimierz Pulaski University of Technology and Humanities in Radom
ul. Stasieckiego54, 26-600 Radom, Poland

E-mails: arturas.kilikevicius@vgtu.lt, vytautas.rimsa@vgtu.lt, m.rucki@uthrad.pl

Article citation info:
HUANG H-Z, YU K, HUANG T, LI H, QIAN H-M. Reliability estimation for momentum wheel bearings considering frictional heat. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 6–14, http://dx.doi.org/10.17531/ein.2020.1.2.

Hong-Zhong HUANG
Kang YU
Tudi HUANG
He LI
Hua-Ming QIAN

RELIABILITY ESTIMATION FOR MOMENTUM WHEEL BEARINGS CONSIDERING FRICTIONAL HEAT

OCENA NIEZAWODNOŚCI ŁOŻYSK KÓŁ ZAMACHOWYCH Z UWZGLĘDNIENIEM CIEPŁA TARCIA

Momentum wheels are the key components of the inertial actuators in the satellites, and the momentum wheel bearings are weak links of momentum wheels as they operate under harsh conditions. The reliability estimation for momentum wheel bearings are helpful to guarantee the mission successes for both momentum wheels and satellites. Hence, this paper put emphasis into reliability estimation of a momentum wheel bearing considering multiple coupling operating conditions and frictional heat by using the finite element analysis. The stress-strength interference model is employed to calculate the reliability of the momentum wheel bearing. A comparative analysis for reliability estimation with and without frictional heat of the momentum wheel bearing is conducted. The results show that the frictional heat cannot be ignored in the reliability analysis of momentum wheel bearings.

Keywords: *Momentum wheel bearings, Finite element analysis, Frictional heat, Stress-strength interference model.*

Koła zamachowe są kluczowymi elementami składowymi siłowników bezwładnościowych w satelitach. Ich łożyska stanowią słaby punkt podczas pracy w trudnych warunkach. Ocena niezawodności łożysk kół zamachowych pozwala zapewnić powodzenie misji zarówno w odniesieniu do samych kół zamachowych, jak i satelitów. Dlatego też niniejszy artykuł poświęcono zagadnieniu oceny niezawodności łożyska koła zamachowego z wykorzystaniem analizy metodą elementów skończonych przy uwzględnieniu wielu sprzężonych warunków pracy oraz ciepła tarcia. Do obliczenia niezawodności łożyska koła zamachowego zastosowano model obciążeniowo-wytrzymałościowy. Przeprowadzono także analizę porównawczą oceny niezawodności łożyska koła zamachowego z uwzględnieniem lub bez uwzględnienia ciepła tarcia. Wyniki pokazują, że w analizie niezawodności łożysk kół zamachowych nie można pominąć ciepła tarcia.

Słowa kluczowe: *łożyska kół zamachowych, analiza MES, ciepło tarcia, model obciążeniowo-wytrzymałościowy.*

Acronyms and Abbreviations

FEA	Finite element analysis
MW	Momentum wheel
MWB	Momentum wheel bearing
PDF	Probability distribution function
SSI	Stress-strength interference

Notations

$N(\cdot)$	Normal distribution
$P(\cdot)$	Probability for a random variable
$R(\cdot)$	Reliability
s	Stress
σ	Strength
$f_s(x)$	PDF of the stress
$f_\sigma(x)$	PDF of the strength
$\Phi(\cdot)$	Standard normal cumulative distribution function

1. Introduction

Momentum wheels (MWs) are used for changing orientation of satellites. With advantages of low-power consumption, strong anti-disturbance ability, contamination-free and high-control precision, MWs have been widely adopted in attitude and orbit control system of satellites. However, the failure of MWs frequently occurred, according to the observation data from [33]. On the one hand, the failure of MWs will lead to catastrophic consequences to inertial actuators, even to the whole satellites [5]. On the other hand, the attitude control of satellites relies on MWs. Hence, the reliability estimation for MWs is essential to guarantee high reliability and mission success of satellites [4].

Due to the limitations of cost, time and difficulties in data collection, reliability estimation of MWs is always restricted by small amount data [22, 18, 21, 35]. To this end, the reliability estimation of MWs mainly relies on the simulation, testing and degradation models [3, 12, 19, 27, 28, 30]. The Russian Academy of Science adopted degradation based method and related testing technology to estimate the lifetime of the turned gyroscope RBHK05-78, and a remaining useful life of 30000 hours was estimated by using the extrapolation method [10]. The venerable companies, BENDIX of USA and TELDIX of German, led the research of MWs. To test the reliability of permanent

magnetic rate integrating gyroscopes, BENDIX took 326 products as samples and spent 794240 hours. TELDIX has provided 588 momentum wheels for 235 satellites, and the total no-failure on-orbit time is up to 2200 years [13]. Jin [9] proposed a method based on the relationship between the MWs' physical performances and their failure mechanisms to evaluate the reliability of MWs. In this method, a stochastic threshold Gauss-Brown process model was put forward to describe the failure process of MWs. Liu [23] and Jin [8] introduced Bayesian method to fuse multi-source information to carry out reliability assessment of MWs. According to engineering practice and testing results, the reliability of MWs relies on the high reliability and long usage life of momentum wheel bearings (MWBs), which are weak links of MWs. From field collection point of view, failures of MWBs occupied the first place in the failures of MWs [31, 32]. The methods of estimating the reliability of WMs often cost a lot of time and money currently. Though reliability estimation on bearings is widely conducted under certain working condition [6, 7, 11, 14, 15, 21, 29, 37, 38], the method of estimating the reliability of MWBs under working condition based finite element analysis (FEA) is rare. Furthermore, this paper takes frictional heat into consideration by adopting the FEA and stress-strength interference (SSI) model to evaluate the reliability of the target MWB. A comparative analysis was conducted, which indicates that the reliability of the MW is relatively affected by the frictional heat. A method based on FEA and SSI model is introduced here, which provides a solution to the reliability estimation for the small sample size products.

The rest of this paper is organized as follows. The finite element analysis of the MWB is carried out in section 2. Section 3 introduces the SSI model. SSI based reliability estimation for the MWB of the MW is conducted in Section 4. Section 5 provides the conclusions of this research.

2. Finite Element Analysis

2.1. Introduction of the MWB

This paper analyzes a MWB of a real MW equipped in a satellite of China. The MWB is a deep groove ball bearing with an inner ring, an outer ring, eight rolling elements and a cage.

The software namely SOLIDWORKS is used to build the three dimensional model of the MWB, and ANSYS WORKBENCH is employed to carry out FEA. The three dimensional model is shown in Fig. 1 and each MWB consists of inner ring, outer ring, cage and 8 rolling elements. The main parameters of the MWB are listed in Table 1.

Table 1. The main parameters of the MWB

Parameter	value	Parameter	value
Bearing outside diameter (mm)	50	Number of balls	8
Bearing bore diameter (mm)	22	Rib diameter of outer ring (mm)	40
Bearing width (mm)	14	Rib diameter of inner ring (mm)	32
Ball diameter (mm)	7		

The MWB was made by GCr15, which has great performance on corrosion resistance and high strength. The density (7830kg/m3), Poisson's ratio (0.3) and coefficient of linear expansion (1.25e-5/°C) of GCr15 are regarded as constant in this study. The changes of elasticity modulus and thermal conductivity of GCr15 together with temperature are shown in Table 2.

Table 2. Material properties of GCr15 changing along with time

Temperature T/°C	Elasticity modulus E/GPa	Thermal conductivity $k/W/(m{\cdot}°C)$
20	200	40.1
200	192	37.85
400	175	34.5
600	153	30.1

2.2. Static-structure Analysis

Contact pairs are automatically generated when the model of MWB is imported into the ANSYS WORKBENCH. The model of the MWB has 8 rolling elements and 24 pairs of contact. Taking the surfaces of cage holes and raceway's groove surfaces of the inner ring and the outer ring as target surfaces, and taking the sphere surfaces of balls as contact surfaces, 24 pairs in total of contact are built in the analysis. All of the contact types are defined as frictional contact with the friction coefficient 0.02.

Meshing of the finite element analysis affects the results directly. Meshing is related to both accuracy of analysis results and calculation time. This paper takes these both parameters into consideration. Moreover, a module named Mesh Quality Check is applied to ensure the quality of the meshing element. The meshing of the MWB creates 10084 elements and 29739 nodes, as shown in Fig. 2.

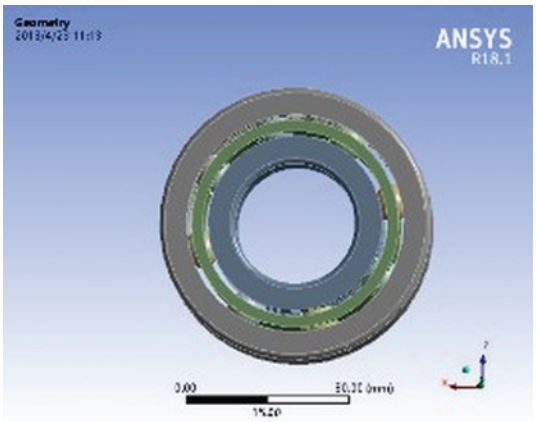


Fig. 1. Three dimensional model of the MWB

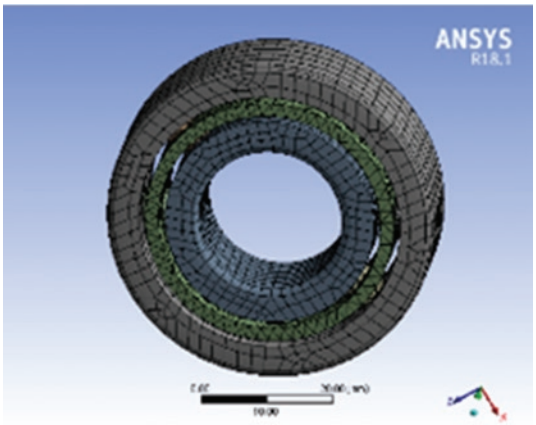


Fig. 2. The meshing of the bearing

Loads and boundary conditions

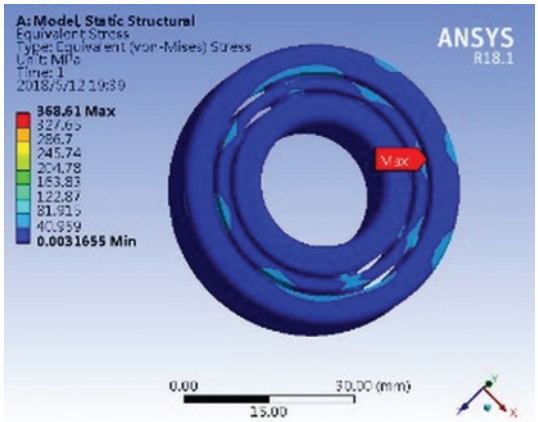
The inner ring of MWB is fixed to the shaft while the outer ring revolves around the axis.

1) To simulate the assemblage of the inner ring and the shaft, a fixed support constraint is applied to the inner surface of the inner ring.
2) To simulate the role of the sleeve bearing, axial displacement constraints are applied to end surfaces of the outer ring, the inner ring and the cage.
3) To simulate the rotation and the assemblage of the outer ring, a rotational velocity is applied to the whole outer ring around the axial orientation, and a radial force, which pointes into inside, is applied to the outside surface of the outer ring.

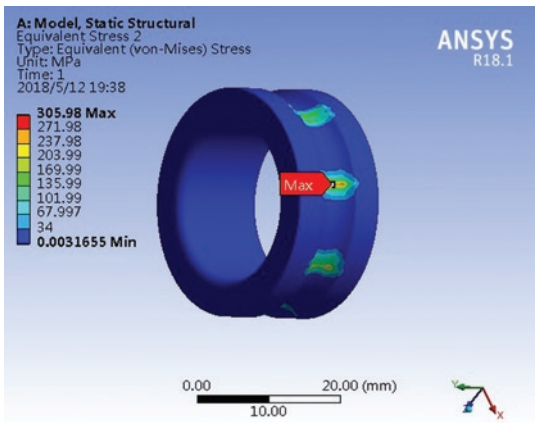
Static structure analysis result

In static structure analysis module, equivalent stress is added in the solution to detect the stress of the riskiest point on the MWB. Fig. 3 shows the results of the equivalent stress of the MWB when the rotation velocity of the outer ring is 10000 rpm, and the radial force of the outer ring is 1200 N. The maximum stress points of the inner ring, outer ring and cage of the appeared at the position of contact with the rolling elements. The maximum stress on different parts of the MWB are listed in Table 3.

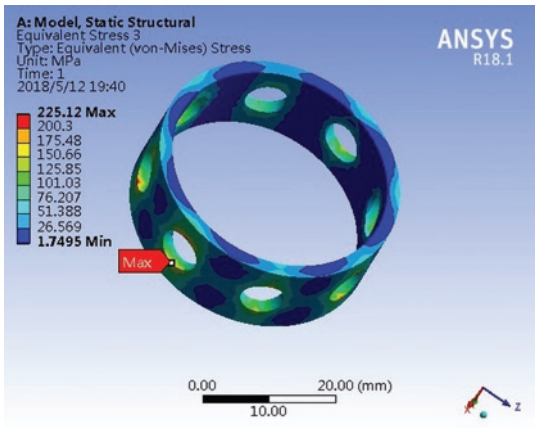
As shown in Table 3 and Fig. 3, the maximum stress is on the rolling element, more specifically, the contact region of the ball and the outer ring. The maximum stress is 368.61MPa, which is regarded as the riskiest point. The second maximum stress part of the MWB is the inner ring, amounting to 305.98MPa. The third maximum part is



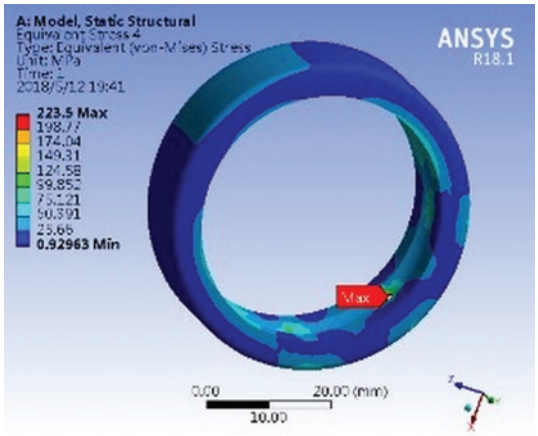
(a) The whole bearing



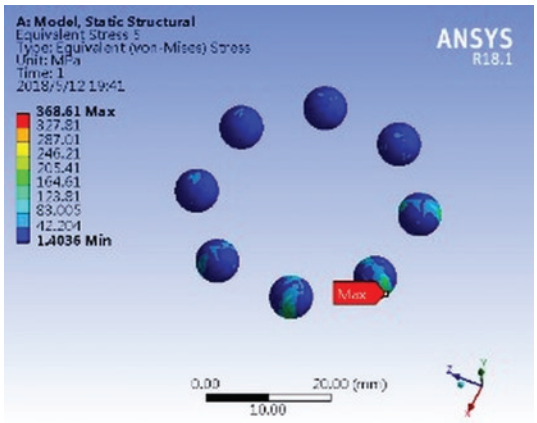
(b) The inner ring



(c) The cage



(d) The outer ring



(e) The rolling elements

Fig. 3. Static structure analysis results

Table 3. The maximum stress on different parts of the MWB

Different part of the MWB	The maximum stress (MPA)	Different part of the MWB	The maximum stress (MPA)
The inner ring	305.98	The outer ring	225.12
The cage	223.50	The rolling elements	368.61

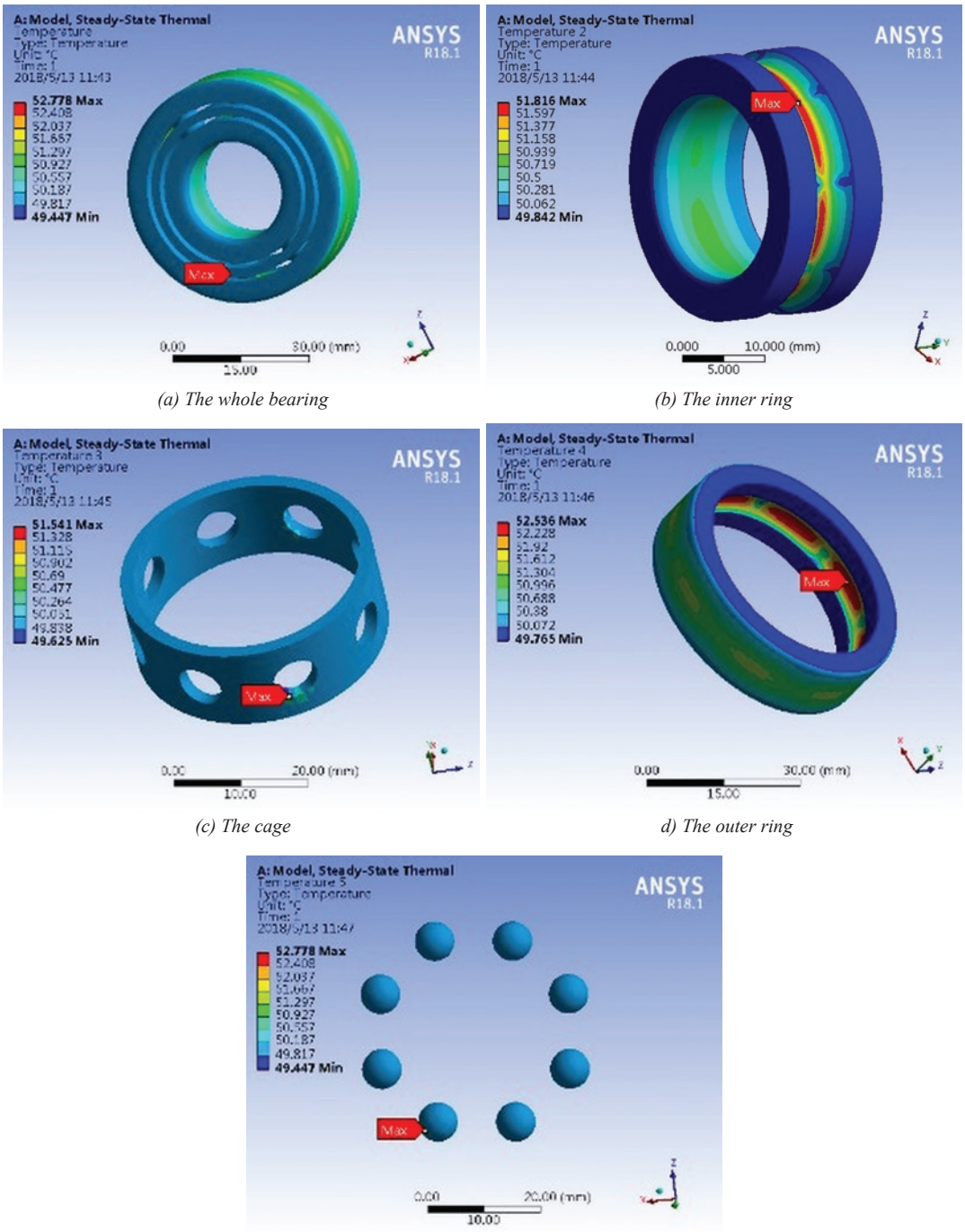
the cage, amounting to 225.12MPa. The minimum stress part is the outer ring, amounting to 223.5MPa. The result in Table 3 is accordance with factual reality, which indicates the correctness of the FEA in this paper.

2.3. FEA of the MWB Considering Frictional Heat

The MWB produces frictional heat between the rolling elements and inner ring, the outer ring and cage. The friction may influence the stress of the riskiest point of the MWB. To this end, this paper adopted indirect coupling method in FEA. Steady-state thermal analysis is firstly completed, then the result is imported to static structure analysis as a thermal load. Finally, the thermal-mechanical FEA is accomplished.

Steady-state thermal analysis

(1) According to the installation and working conditions of MWB, the Loads and boundary conditions are shown as follows:



(a) The whole bearing

(b) The inner ring

(c) The cage

d) The outer ring

(e) The rolling elements

Fig. 4. The steady-state thermal analysis results

1) The heat generated by the friction between the rolling elements and the raceways is loaded in the form of heat flux to the raceway surfaces, which is 42W.
2) The heat generated by the friction between the rolling elements and the cage is applied in the form of heat flux to the hole surfaces of the cage, which is 5W.
3) Applied the convection to the outer surfaces of the inner ring, outer ring and the body surface, which is 60 $W/(mm^2 \cdot °C)$.

(2) The results of steady-state thermal analysis of different parts of the WMB are shown in Fig. 4.

From Fig. 4, the highest temperature of the inner ring, outer ring and retainer appears at the position of contact with the rolling elements. And the highest temperature is 52.778°C, which locates on the rolling element, specifically, the contact area between the ball and the outer ring. The second highest temperature part of the MWB is the outer ring, amounting to 52.536°C. The third highest temperature part

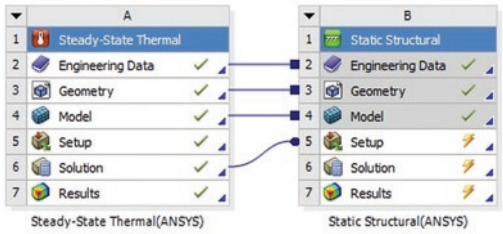
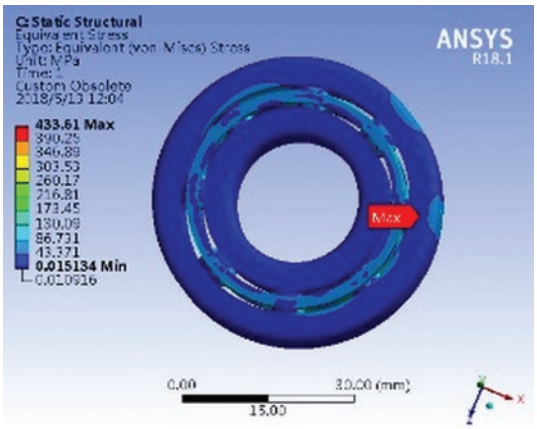
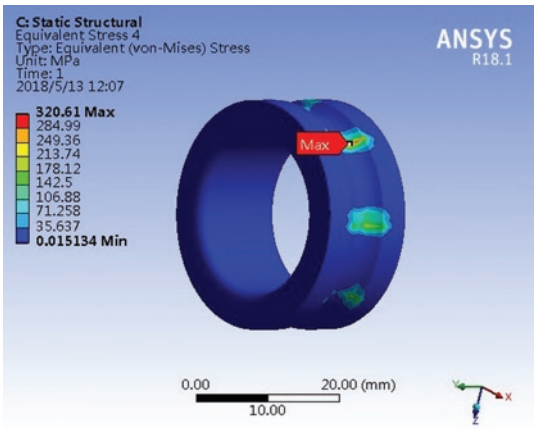


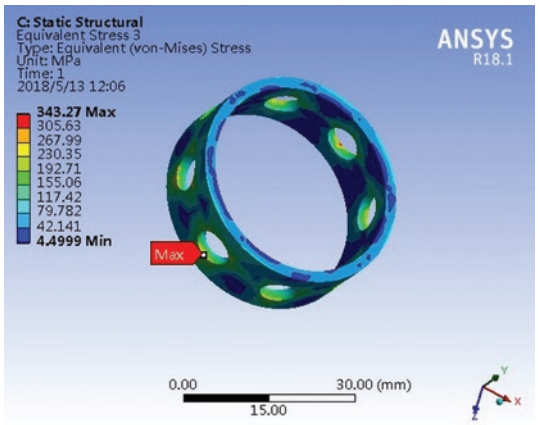
Fig. 5. The process of thermal-mechanical coupling analysis



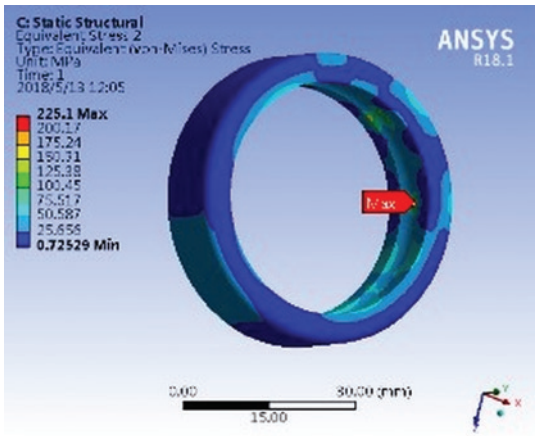
(a) The whole bearing



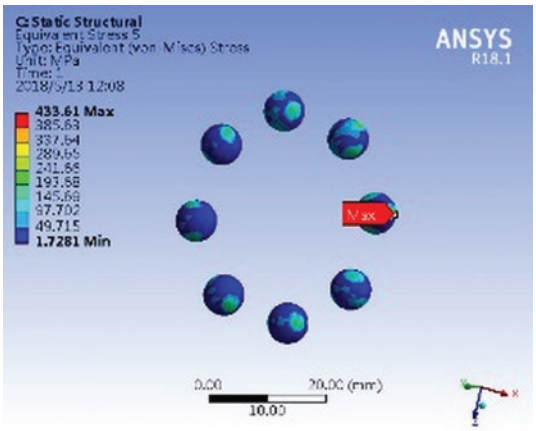
(b) The inner ring



(c) The cage



(d) The outer ring



(e) The rolling elements

Fig. 6. The results of thermal-mechanical coupling FEA

of the MWB is inner ring, amounting to 51.816°C. The fourth highest temperature part of the MWB is cage, amounting to 51.541°C.

Thermal-mechanical coupling analysis

This paper considers both efficiency of calculation and working conditions of the MWB. The process is shown in Fig. 5.

The result of steady-state thermal analysis is imported into the static-structure analysis as the thermal load. Other loads and boundary conditions are same to the static-structure analysis. The results of thermal-mechanical coupling FEA are shown in Fig. 6.

It can be seen from Fig. 6 that the maximum stress is up to 433.61MPa located on the contact area between the rolling elements and the outer ring. It is evident that the stress on the WMB increases considerably due to the frictional heat, which indicates that the thermal field cannot be ignored when undertaking FEA to the MWB. That is to say the reliability estimation result is more accurate when taking frictional heat into consideration than not. It is shown in Fig. 6 that the maximum stress is on the rolling element, more specifically, the contact region of the ball and the outer ring. The maximum stress is 433.61MPa, which is regarded as the riskiest point.

3. Stress-Strength Interference Model

Stress-strength interference (SSI) model has been widely used in the reliability estimation [1, 2, 25, 36]. In SSI model, Stress represents a number of factors promoting the failure while strength represents ability of resisting the failure of products. Products fail or not depends on the relationship between the stress and the strength. Failure occurs if the stress is larger than the strength, otherwise structures would be safe [16, 17, 20]. In practice, uncertainties exist in both stress and strength, thus stress and strength is no longer a constant value as most researchers did [34]. It is proper to describe them by distributions [24]. The reliability of a structure is defined as the probability of strength is larger than stress.

The probability density function (PDF) of strength is represented as $f_\delta(y)$, and the PDF of stress is represented as $f_s(x)$.According to the definition of the SSI model, the reliability of the structure is calculated by Eq. (1):

$$R = P(Y - X > 0) = \int_{-\infty}^{\infty} f_\delta(y)\left[\int_{-\infty}^{Y} f_s(x)dx\right]dy \qquad (1)$$

If both X and Y follow normal distribution $X \sim N(\mu_x, \sigma_y^2)$, $Y \sim N(\mu_y, \sigma_y^2)$, their proper function is $Z = Y - X$. According to the property of normal distribution, Z is also follow a normal distribution $Z \sim N(\mu_z, \sigma_z^2)$. Moreover, $\mu_z = \mu_y - \mu_x$, $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$. Then, Eq. (1) can be transformed to Eq. (2):

$$\begin{aligned} R &= P(Y - X > 0) \\ &= P(Z > 0) \\ &= \int_0^{\infty} \frac{1}{\sqrt{2\pi}\, s_z} \exp\left[-\frac{1}{2}\left(\frac{Z-\mu_z}{\sigma_z}\right)^2\right] dZ \\ &= \Phi\left(\frac{Z-\mu_z}{\sigma_z}\right) \end{aligned} \qquad (2)$$

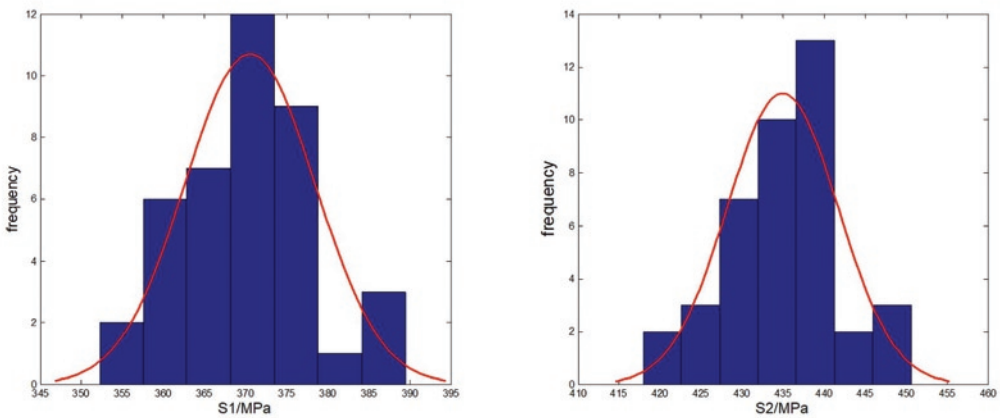
4. Reliability Estimation of MWB

Owning to complex working conditions, the rotational velocity of the MWB is not a constant value but a distribution. The rotational speed of the bearing can be represented by a normal distribution with mean value of 10000 and standard deviation of 100, which means $\omega \sim N(10000, 100^2)$ [7]. Taking 40 random numbers from the normal distribution of the rotational speed, and adopted them into FEA. In order to analyze the effect of frictional heat on the reliability of the MWB, both the static structure analysis and thermal-mechanical coupling analysis were undertaken 40 times under different rotational velocity. The stress on the riskiest point of the MWB were calculated by the FEA, as shown in Table 4.

In Table 4, *n* represents the rotational velocity of the outer ring, S_1 and S_2 represents the maximum stress of the bearing under the static structure analysis and thermal-mechanical analysis, respectively. The histograms of stress on the riskiest point of the MWB are shown in Fig. 7.

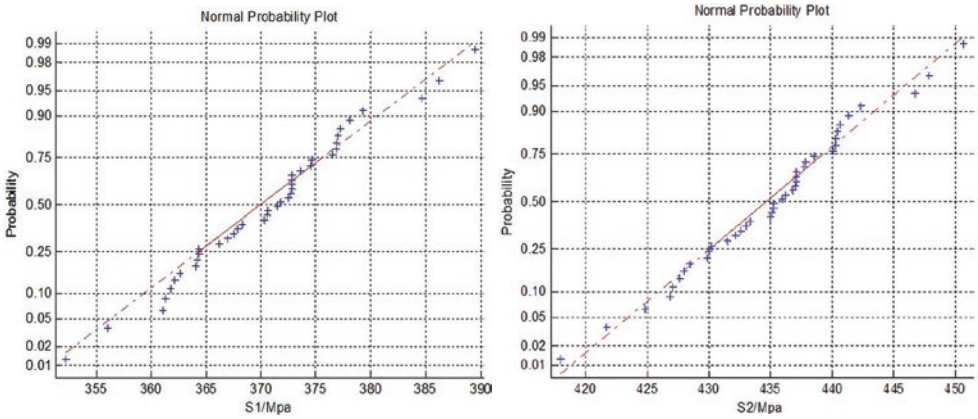
S1 and S2 follow normal distributions, as presented in Fig. 8.

It is can be seen from Fig. 7, there are 40 simulation data in each of the two conditions. Therefore, the S1 and S2 can be fit by normal distribution. Then, Lilliefors test is taken, and the results show that S_1 and S_2 follow normal distribution well. The software of MATLAB was used to fit S_1 and S_2, and the results are $S_1 \sim N(370.6678, 7.7655^2)$ and $S_2 \sim N(434.9405, 6.7781^2)$.



(a) The histogram of S1 *(b) The histogram of S2*

Fig.7. The histograms of stress on the riskiest point of the MWB



(a) The normal plot of S1 *(b) The normal plot of S2*

Fig. 8. The normal plot of the stress on the riskiest point of the MWB

Table 4. The stress on the riskiest point of the MWB at various rotational speed

n/rpm	S_1/MPa	S_2/MPa	n/rpm	S_1/MPa	S_2/MPa
10054	371.79	436.23	10072	372.78	437.06
10183	379.23	442.31	10163	378.07	441.34
9774	355.99	421.71	10049	371.45	435.96
10086	373.60	437.73	10103	374.59	438.50
10032	374.47	435.13	10073	372.83	437.10
9869	361.27	427.06	9970	366.94	432.15
9957	366.19	431.51	10029	370.30	434.99
10034	370.63	435.26	9921	364.15	429.97
10358	389.45	450.63	10089	372.82	437.87
10277	384.67	446.74	9885	362.17	427.97
9865	361.06	426.84	9893	362.60	428.40
10303	386.19	447.86	9919	364.04	429.86
10073	372.83	437.10	9706	352.25	417.96
9994	368.28	433.33	10144	376.96	440.42
10071	372.72	437.01	10033	370.57	435.22
9980	367.48	432.61	9925	364.36	430.18
9988	367.91	433.03	10137	376.52	440.08
10149	377.24	440.65	9829	359.05	424.80
10141	376.80	440.31	10067	372.50	436.83
10142	376.85	440.32	9879	361.80	427.59

Due to insufficient data of MWB, and the influence of manufacturing technology and assembly level, it is difficult to obtain the intensity distribution of MWB. In this paper, the intensity distribution of material is approximately used as the intensity distribution of MWB. The strength of GCr15 follows normal distribution $\delta \sim N\left(518.42, 51.842^2\right)$.

According to the SSI model, the reliability of the MWB can be calculated by Eq. (2):

$$R_1 = P(\delta - S_1 > 0) = \Phi\left(\frac{518.42 - 370.6678}{\sqrt{51.842^2 + 7.7655^2}}\right) = 0.99759 \qquad R_2 = P(\delta - S_2 > 0) = \Phi\left(\frac{518.42 - 434.9405}{\sqrt{51.842^2 + 6.7781^2}}\right) = 0.94483$$

R_1 and R_2 represent the reliability of the MWB under the static-structure analysis and under the thermal-mechanical coupling analysis, respectively. The results indicate that frictional heat does affect the reliability of MWB, and frictional heat cannot be ignored in the reliability estimation of the MWB.

5. Conclusions

In this paper, a reliability estimation method for MWBs based on FEA and SSI model is proposed. Firstly, the thermal-mechanical coupling FEA to the MWB is carried out. Then, the reliability estimation based on stress-strength interference model and the comparison analysis between the static-structure analysis and the thermal-mechanical coupling analysis is fulfilled. The reliability under static-structure analysis is 0.99759, and the reliability under the thermal-mechanical

analysis is 0.94483. It is evident that the reliability decreases a lot owing to the effect of frictional heat, which indicates that the effect of frictional heat on the reliability of MWB cannot be ignored. Put it in another way, the reliability estimation result considering frictional heat proposed in this paper is more accurate than the method ignoring the fictional heat. Future work can be done by considering more physical fields, for example, electrical field, magnetic field and so on to make the estimation result more close to practical engineering.

Acknowledgement

This work was supported by the National Natural Science Foundation of China under contract No. 51875089.

References

1. Al-Mutairi D K, Ghitany M E, Kundu D. Inferences on stress-strength reliability from Lindley distributions. Communications in Statistics-Theory and Methods 2013; 42(8): 1443-1463, https://doi.org/10.1080/03610926.2011.563011.
2. Bhuyan P, Dewanji A. Reliability computation under dynamic stress-strength modeling with cumulative stress and strength degradation. Communications in Statistics-Simulation and Computation 2017; 46(4): 2701-2713, https://doi.org/10.1080/03610918.2015.1057288.
3. Bingjie W, Guang J. Reliability modeling and analyzing of momentum wheel based on Gamma process. Value Engineering 2010; (1): 25-27.
4. Castet J F, Saleh J H. Beyond reliability, multi-state failure analysis of satellite subsystems: a statistical approach. Reliability Engineering & System Safety 2010; 95(4): 311-322, https://doi.org/10.1016/j.ress.2009.11.001.
5. Castet J F, Saleh J H. Satellite and satellite subsystems reliability: Statistical data analysis and modeling. Reliability Engineering & System Safety 2009; 94(11): 1718-1728, https://doi.org/10.1016/j.ress.2009.05.004.
6. Huang C G, Huang H Z, Li Y F. A Bi-Directional LSTM prognostics method under multiple operational conditions. IEEE Transactions on Industrial Electronics 2019; 66(11): 8792-8802, https://doi.org/10.1109/TIE.2019.2891463.
7. Jansen M J, Jones Jr W R, Pepper S V, Wheeler D R, Schroeer A, Fluehmann F, Shogrin B A. The effect of TiC coated balls and stress on the lubricant lifetime of a synthetic hydrocarbon (Pennzane 2001A) using a vacuum spiral orbit tribometer. In Proceeding International Tribology Conference, Nagasaki, Japan; October 2000.
8. Jin G, Feng J. Bayes-Weibull reliability assessment method for long life satellite moving components. Systems Engineering and Electronics 2009; 31(8): 2020-2023.
9. Jin G, Liu Q, Zhou J, Zhou Z. Repofe: Reliability physics of failure estimation based on stochastic performance degradation for the momentum wheel. Engineering Failure Analysis 2012; 22: 50-63, https://doi.org/10.1016/j.engfailanal.2011.12.004.
10. Jin G, Matthews D, Fan Y, Liu Q. Physics of failure-based degradation modeling and lifetime prediction of the momentum wheel in a dynamic covariate environment. Engineering Failure Analysis 2013; 28: 222-240, https://doi.org/10.1016/j.engfailanal.2012.10.027.
11. Khonsari M, Booser E R. Predicting lube life-heat and contaminants are the biggest enemies of bearing grease and oil. Machinery Lubrication 2003; 75(1): 89-90, 92.
12. Li H, Huang H Z, Li Y F, Zhou J, Mi J. Physics of failure-based reliability prediction of turbine blades using multi-source information fusion. Applied Soft Computing 2018; 72: 624-635, https://doi.org/10.1016/j.asoc.2018.05.015.
13. Li H, Pan D, Chen C L P. Reliability modeling and life estimation using an expectation maximization based wiener degradation model for momentum wheels. IEEE transactions on cybernetics 2014; 45(5): 969-977, https://doi.org/10.1109/TCYB.2014.2341113.
14. Li X Y, Huang H Z, Li Y F, Zio E. Reliability assessment of multi-state phased mission system with non-repairable multi-state components. Applied Mathematical Modelling 2018; 61: 181-199, https://doi.org/10.1016/j.apm.2018.04.008.
15. Li X Y, Li Y F, Huang H Z, Zio E. Reliability assessment of phased-mission systems under random shocks. Reliability Engineering & System Safety 2018; 180: 352-361, https://doi.org/10.1016/j.ress.2018.08.002.
16. Li X, Huang H Z, Li F, Ren L. Remaining useful life prediction model of the space station. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(3): 501-510, https://doi.org/10.17531/ein.2019.3.17.
17. Li X, Huang H Z, Li Y F, Li Y F. Reliability evaluation for VHF and UHF bands under different scenarios via propagation loss model. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(3): 375-383, https://doi.org/10.17531/ein.2019.3.3.
18. Li Y F, Huang H Z, Liu Y, Xiao N, Li H. A new fault tree analysis method: fuzzy dynamic fault tree analysis. Eksploatacja i Niezawodnos c-Maintenance and Reliability 2012; 14(3): 208-214.
19. Li Y F, Huang H Z, Mi J, Peng W, Han X. Reliability analysis of multi-state systems with common cause failures based on Bayesian network and fuzzy probability. Annals of Operations Research 2019; https://doi.org/10.1007/s10479-019-03247-6, https://doi.org/10.1007/s10479-019-03247-6.
20. Li Y F, Mi J, Huang H Z, Xiao N C, Zhu S P. System reliability modeling and assessment for solar array drive assembly based on Bayesian networks. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(2): 117-122.
21. Li Y F, Mi J, Huang H Z, Zhu S P, Xiao N. Fault tree analysis of train rear-end collision accident considering common cause failure. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(4): 403-408.
22. Liu Q, Jin G, Zhou J. A modeling method of performance reliability of momentum wheel based on EMD. Computer Simulation 2007; 24(11): 32-34, 158.
23. Liu Q, Zhou J, Jin G, Li H T. Bayesian reliability estimation for momentum wheel based on credibility. Journal of Astronautics 2009; 30(1): 382-386.
24. Liu Y, Shi Y, Bai X, Liu B. Stress-strength reliability analysis of system with multiple types of components using survival signature. Journal of Computational and Applied Mathematics 2018; 342: 375-398, https://doi.org/10.1016/j.cam.2018.04.029.
25. Liu Y, Shi Y, Bai X, Zhan P. Reliability estimation of a NM-cold-standby redundancy system in a multicomponent stress-strength model with generalized half-logistic distribution. Physica A: Statistical Mechanics and its Applications 2018; 490: 231-249, https://doi.org/10.1016/j.physa.2017.08.028.
26. Masuko M, Mizuno H, Suzuki A, Obara S, Sasaki A. Lubrication performance of multialkylatedcyclopentane oils for sliding friction of steel under vacuum condition. Journal of Synthetic Lubrication 2007; 24(4): 217-226, https://doi.org/10.1002/jsl.41.

27. Mi J, Li Y F, Peng W, Huang H Z. Reliability analysis of complex multi-state system with common cause failure based on evidential networks. Reliability Engineering & System Safety 2018; 174: 71-81, https://doi.org/10.1016/j.ress.2018.02.021.
28. Mi J, Li Y F, Yang Y J, Peng W, Huang H Z. Reliability assessment of complex electromechanical systems under epistemic uncertainty. Reliability Engineering & System Safety 2016; 152: 1-15, https://doi.org/10.1016/j.ress.2016.02.003.
29. Palladino M, Murer J, Didierjean S, Gaillard L. Life prediction of fluid lubricated space bearings: A case study. In Proc. 14th Eur. Space Mechanisms Tribol. Symp 2011; 279-285.
30. Prado J, Bisiacchi G, Reyes L, Vicente E, Contreras F, Mesinas M, Juares A. Three-axis air-bearing based platform for small satellite attitude determination and control simulation. Journal of Applied Research and Technology 2005; 3(3): 222-237.
31. Sathyan K, Gopinath K, Lee S H, Hsu H Y. Bearing retainer designs and retainer instability failures in spacecraft moving mechanical systems. Tribology Transactions 2012; 55(4): 503-511, https://doi.org/10.1080/10402004.2012.675118.
32. Sathyan K, Hsu H Y, Lee S H, Gopinath K. Long-term lubrication of momentum wheels used in spacecrafts-an overview. Tribology International 2010; 43(1-2): 259-267, https://doi.org/10.1016/j.triboint.2009.05.033.
33. Tafazoli M. A study of on-orbit spacecraft failures. Acta Astronautica 2009; 64(2-3): 195-205, https://doi.org/10.1016/j.actaastro.2008.07.019.
34. Wang B X, Geng Y, Zhou J X. Inference for the generalized exponential stress-strength model. Applied Mathematical Modelling 2018; 53: 267-275, https://doi.org/10.1016/j.apm.2017.09.012.
35. Xu H, Li W, Li M, Hu C, Zhang S, Wang X. Multidisciplinary robust design optimization based on time-varying sensitivity analysis. Journal of Mechanical Science and Technology 2018; 32(3): 1195-1207, https://doi.org/10.1007/s12206-018-0223-8.
36. Zhang J, Ma X, Zhao Y. A stress-strength time-varying correlation interference model for structural reliability analysis using copulas. IEEE Transactions on Reliability 2017; 66(2): 351-365, https://doi.org/10.1109/TR.2017.2694459.
37. Zhang X, Gao H, Huang H Z, Li Y F, Mi J. Dynamic reliability modeling for system analysis under complex load. Reliability Engineering & System Safety 2018; 180: 345-351, https://doi.org/10.1016/j.ress.2018.07.025.
38. Zheng B, Li Y F, Huang H Z. Intelligent fault recognition strategy based on adaptive optimized multiple centers. Mechanical Systems and Signal Processing 2018; 106: 526-536, https://doi.org/10.1016/j.ymssp.2017.12.026.

Hong-Zhong HUANG
Kang YU
Tudi HUANG
He LI
Hua-Ming QIAN

Center for System Reliability and Safety
School of Mechanical and Electrical Engineering
University of Electronic Science and Technology of China
Chengdu, Sichuan, 611731, P. R. China

E-mails: hzhuang@uestc.edu.cn, yukang@std.uestc.edu.cn, huangtudi@std.uestc.edu.cn, heli@std.uestc.edu.cn, qianhuaming@std.uestc.edu.cn

Article citation info:
ANDRUSZKIEWICZ J, LORENC J, ŁOWCZOWSKI K, WEYCHAN A, ZAWODNIAK J. Energy losses' reduction in metallic screens of MV cable power lines and busbar bridges composed of single-core cables. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 15–25, http://dx.doi.org/10.17531/ein.2020.1.3.

Jerzy ANDRUSZKIEWICZ
Józef LORENC
Krzysztof ŁOWCZOWSKI
Agnieszka WEYCHAN
Józef ZAWODNIAK

ENERGY LOSSES' REDUCTION IN METALLIC SCREENS OF MV CABLE POWER LINES AND BUSBAR BRIDGES COMPOSED OF SINGLE-CORE CABLES

OGRANICZANIE STRAT ENERGII W ŻYŁACH POWROTNYCH LINII I MOSTÓW KABLOWYCH ŚREDNIEGO NAPIĘCIA WYKORZYSTUJĄCYCH KABLE JEDNOŻYŁOWE*

The growing share of medium voltage cable lines in distribution networks challenges distribution network operators in terms of proper mode of operation of these lines. It is related to the reduction of energy losses in cable conductors and metallic cable screens. The article focuses on energy losses in metallic cable screens of cable lines and substation busbar bridges composed of single-core cables with metallic screens and possible ways of their reduction. Simulation and measurement analysis of the level of energy losses in the metallic screens of cables is presented together with the economic analysis of various variants of losses reduction through the change of the way these screens are operated in relation to the traditional bilateral earthing at both ends of cable. Technical problems and threats connected with the use of considered modifications of metallic screens operation during earth fault disturbances in distribution networks are also presented.

***Keywords**: cable lines, medium voltage network, energy losses, metallic cable screen.*

Rosnące skablowanie linii średniego napięcia w sieciach dystrybucyjnych stawia przed operatorami tych sieci wyzwanie prawidłowej eksploatacji linii kablowych. Powiązane jest to z redukowaniem strat energii w żyłach roboczych i powrotnych kabli. W artykule skupiono się na stratach energii w żyłach powrotnych linii oraz mostów kablowych wykonanych przy wykorzystaniu kabli jednożyłowych z metalicznymi żyłami powrotnymi oraz możliwych sposobach ich ograniczania. Przedstawiono analizę symulacyjną i pomiarową poziomu strat energii w żyłach powrotnych kabli wraz z analizą ekonomiczną różnych wariantów ich redukcji poprzez zmianę sposobu pracy tych żył w stosunku do tradycyjnego obustronnego ich uziemienia. Przedstawione zostały również problemy techniczne oraz zagrożenia związane z zastosowaniem rozważanych modyfikacji pracy żył powrotnych podczas zakłóceń zwarciowych w sieciach dystrybucyjnych.

***Słowa kluczowe**: linie kablowe, sieć średniego napięcia, straty energii, żyła powrotna.*

1. Introduction

SAIDI (System Average Interruption Duration Index), which is characterizing the reliability of electricity supplies from the medium voltage distribution network, is significantly greater in Poland than in other European countries, which are the leaders in these statistics. One of the important reasons for this state is the relatively low share of cable lines in medium voltage (MV) network, not exceeding 25% in Poland [4]. The expected profits from converting the MV network into the cable one can be assessed by comparing average SAIDI values in Polish distribution network and in foreign networks having the high share of cable lines. The SAIDI of distribution networks in Switzerland, Denmark, Luxembourg, Germany and the Netherlands, basically operating a cable network, reach the level several times lower than in the case of networks with a significant share of overhead power lines in the MV network, as illustrated in Fig. 1.

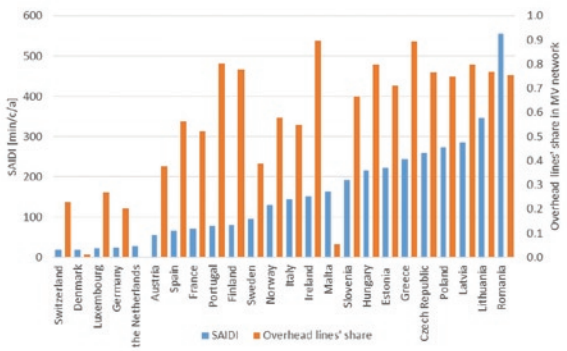


Fig. 1. SAIDI in comparison with the share of overhead lines in the total length of the medium voltage distribution network in various countries based on [4]

In order to improve this situation in Poland, a broad plan for replacing the overhead lines by cables in the sections of the network characterized by high statistical damage indicators, usually passing the forests or supplying wooded areas, is being prepared. In the coming twenty years, one should therefore expect a

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

significant increase in the MV cable network at the level of 40 000 km [40]. Replacement of existing overhead lines by cable lines will allow to eliminate the impact of severe climatic conditions and increase the operational reliability of these sections [17, 29]. Moreover, cable lines have lower impact on the landscape and are more easily accepted by the public [49].

In the last two decades, there have also been significant changes in cable line technology. Practically, the most common insulating material used is crosslinked polyethylene, and for the construction of cable lines, single-core cables with a metallic screen are used to ensure effective earth fault short-circuit current discharge and enable rapid operation of the damaged cable line protection system to prevent long-term effects of this current on insulation layers of other phases of a cable and its surroundings [37, 43]. In case of the absence of a cable metallic screen in the vicinity of short circuit event, significant electrical risks could occur as well. Detailed requirements in terms of protection against electric shock in MV networks are presented in [14-16].

The presented circumstances prompt to undertake research on the correct operation of the medium voltage distribution network containing an increasing share of cable line sections constructed in the new technology using three single-core cables with metallic screens. The problem presented in this article concerns the energy losses in cable metallic screens resulting from the use of their earthing at both ends and the ways to prevent such losses by modification of the traditionally used mode of metallic screens operation. The advantages and drawbacks of the proposed modifications are discussed as well. Considered issue is presented using computer simulations of cable line operation, which were verified by measurements conducted in existing sections of MV cable lines in the network of one of the Polish distribution system operators. The presented issue is in line with global trends in reducing energy losses in distribution networks [7].

Energy losses in cables are associated with the occurrence of a core conductor, insulation and metallic cable screens or other metallic layers of the cable. A distinction can be made between losses in the core conductor due to its resistance, dielectric losses related to the cable capacitance and insulation parameters as well as losses in the metallic cable screen related to the current flow through these layers and eddy currents [2, 35]. Losses associated with eddy currents are usually much lower in the cable screens than losses related to the current flow induced in the screen's circuit by load currents and do not depend on the circuit arrangement [37].

Energy losses in cable lines depend on the nominal cross-section of the conductor and of the metallic screen, way of laying the cable (flat or trefoil formation), but also to a large extent, on the connection and earthing system of the cables' metallic screens [32, 34, 38, 52]. The greater the distance between single-core cables, regardless of the thickness of individual conductors, the greater losses in these cable systems are observed, hence much greater losses will occur in the case of cables in a flat formation than in a trefoil formation [20, 37]. If there is more than one cable system, for example two three-phase cables side by side, energy losses will also depend on the phase sequence in individual single-phase cables [38].

The most popular in MV networks in Poland is both end bonding and earthing of metallic screens of a cable line section. In such systems, under normal load operation, there are negligibly low voltages in metallic screens, but significant currents induced by currents flowing in the cable core conductors, which cause additional losses in the cable metallic screens and may reduce the nominal current carrying capacity of the cable line. Losses in cable metallic screens, caused by induced currents, depend on the coupling impedances of the core conductors and metallic screens of the cable [52]. For standard medium voltage cables, additional losses in the metallic screens constitute 2-10% of the total losses in the cable [20]. Due to the large number of factors determining the losses in the metallic screens, in the most unfavourable cases, the loss in the metallic screens may be greater than the losses in the core conductor [18]. Reduction of losses occurring in metallic screens is possible by modification of operation mode of cable metallic screens including their earthing system, such as single-point metallic screen earthing applied in the case of high voltage cables [6, 11, 52] associated with the use of surge arresters at the unearthed screen end [20, 30], cable metallic screen intersection and cross-bonding [6, 11, 21, 33] or cross-bonding and transposition of the load carrying wires as well [44], inserting additional resistances or inductances at the place of cable metallic screens earthing or at cable joints' earthing [24, 32, 46] or reducing the cross-section of cable metallic screen, which are very often oversized [3, 27, 28, 50]. However, these solutions are rarely used in medium voltage networks due to the fear of electric shock hazards or overvoltage of screen insulation [11, 20] and expected problems with detecting cable line operation distortions due to the modified metallic cable screens connection systems [10, 23] and the possibility of damages in additionally introduced cable cross bonding joints or boxes [48]. In this study, a new task is undertaken to determine the cost-effectiveness of the proposed measures to reduce losses in metallic cable screens, and this cost-effectiveness is evaluated in real energy market conditions depending on the load of the line and the method used to decrease losses in the metallic screens.

The proposed methods of reducing losses in the metallic cable screens also require analysis during fault states. The analysis of electric shock hazards in cable networks [41] should take into account the influence of the number of MV/LV substations operating in the considered MV grids and their earthing as well as the resistance of the cable layers on the flow of ground fault current. The distribution of short-circuit current in the earthing system is also the subject of analysis in [42], where the influence of metal elements in the ground is taken into account. In [5], the influence of parameters of cable layers and earthing system on the earth fault current distribution and the hazards associated with its flow are analysed. The new approach presented in this paper concerns simulations and site studies of earth fault current distribution in different cases of earthing of the metallic screens in order to identify the risk of electric shock.

An important element of cable analysis in fault cases are also overvoltages, which may occur at the unearthed ends of the cable screens. Overvoltage issues are most often analysed for high voltage lines [6, 20, 48], less frequently for medium voltage lines [19], but their analysis is extremely important to ensure proper condition of metallic cable screens' insulation. In this article, attention was drawn to the possible occurrence of overvoltages when using the proposed methods of reducing losses in the metallic cable screens.

The article addresses this issue in order to develop proposals for the operation of cable lines built in the presented technology allowing for reduction of losses in cable metallic screens in economically justified cases when simultaneously surge protection requirements concerning the screen insulation and electric shock protection requirements at MV/LV transformer substations are met. For this purpose extensive simulation tests of medium voltage networks with cable sections have been conducted as well as tests on models containing sections of real cable lines and measurements on sections of cables operating in the distribution network, which are described below.

2. Cable lines' modelling

The analyzed sections of cable lines operating in the distribution network were modelled using the DIgSILENT PowerFactory software. In the cable lines modelling procedure, firstly a single-core cable is created, based on which a three-phase cable is composed. The input data applies to all conductive, insulating and semi-conductive layers that occur. All geometric parameters defining the cross-section

and data defining the properties of all component materials are also entered.

Using the defined cable type, it is possible to create a simulation model of the cable system used in distribution networks as shown in Fig. 2. Input data considered for this purpose are: the type of single-core cable used to build the cable system, the position of each cable line in relation to the ground level, the position of each single-core cable line in relation to the remaining ones, as well as the cable laying environment (earth, air), frequency, soil resistivity, or the number of parallel cable systems.

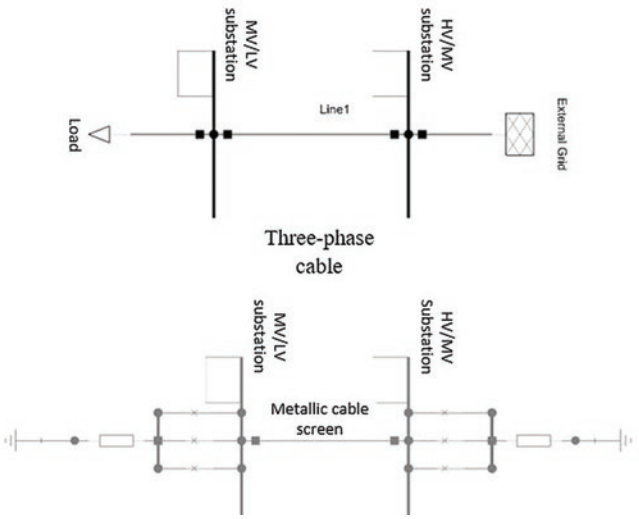


Fig. 2. Cable system model

The calculations for a cable system created with three single-core cables are conducted on the basis of solving matrix differential equations that bind the currents in the cable core conductor and metallic screen with voltages along the conductive layers of the cable [2, 8, 9]. Such calculation method enables the operation of the cable system's simulation in terms of currents and voltages in the conductor and in the metallic screen, in steady and transient conditions during normal operation and short-circuits. The model shown in Fig. 2 allows to simulate the cable operation considering different ways of connecting metallic screens with each other and with the ground and to include various values of earthing electrode resistances. In the case of busbar bridges, a cable system was modelled as consisting of three or four cable bundles connected in a parallel between the HV/MV transformer and the busbars of the MV substation. In the considered case, cables and busbar bridges were modelled in accordance with the MV distribution cable network standard presented in [12].

3. Load and energy losses modelling in MV cable lines

The purpose of this study was to determine the losses in the conductors and metallic screens for the registered loads in selected existing sections of MV cables in order to propose technical solutions to reduce losses in the metallic cable screens. In an urban network, two cable routes were selected which output power from HV/MV substations, consisting of three cable sections connected in series in trefoil formation. For the HV/MV substation bay, to which cable sections were directly connected, actual values of load currents were registered and provided by the Distribution Management System (DMS) controlling the operation of HV/MV substations, for the selected days of 2017 in winter, spring, summer and autumn for periods from Wednesday to Sunday. On the basis of relative changes in energy consumption in particular weeks of the year, the annual load profile for the tested facility was determined. The annual energy demand profile of the analysed cable section created for the maximum load case of 2,0 MVA is shown in Fig. 3.

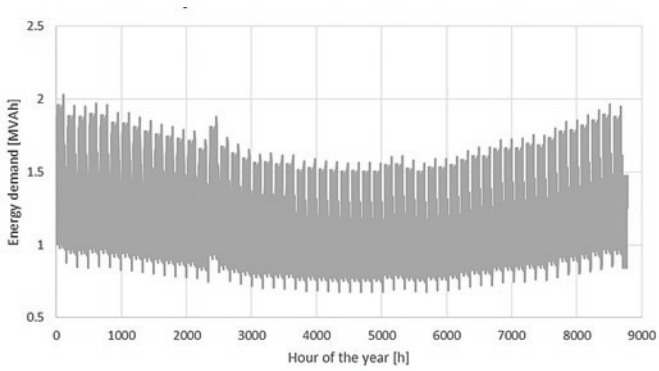


Fig. 3. Annual load profile for the considered cable line bay

The illustrated load profile applies only to the cable section that is connected to the substation bay with the conductor cross-section of 3x240mm^2 and 50 mm^2 for the metallic screens. For the following cable sections, the conductors' cross-section decreased to 3x120 mm^2, due to the reduced load because of additional outflows in the following MV/LV substations, while the same cross-section of the metallic screens was maintained. Taking into account the number and power of MV/LV transformers on a selected line, it is possible to determine the relative load of subsequent cable line sections, assuming an even load of MV/LV distribution transformers proportional to their rated power. In this way, the load on the second section was determined at the level of 72% of the first section load. Similarly, taking into account the next power stations, the third section was loaded at 63% of the first section load. With the progress of balancing meters installation at MV/LV substations, being the component of AMI (Advanced Metering Infrastructure) system, more precise estimation of cable section load will be possible, based on the values registered by the meters, but such data was not available at the time of the presented study.

For the first of the cable lines considered, due to the large differences in the length of the cable sections, the losses in section II and III were much smaller than in section I. Losses in individual sections of the considered cable system indicate seasonal variability, according to the variation of the load modelled in Fig. 3. The conducted analyses allow to determine losses in the metallic screen of the cable, which are presented in the form of a graph in the Fig. 4 as the absolute annual values in kWh and as a share in total energy losses in the cable for the analysed sections. The results of the analysis indicate that the average share of energy losses in metallic screens constitutes 3,5% of the total losses in the cable line.

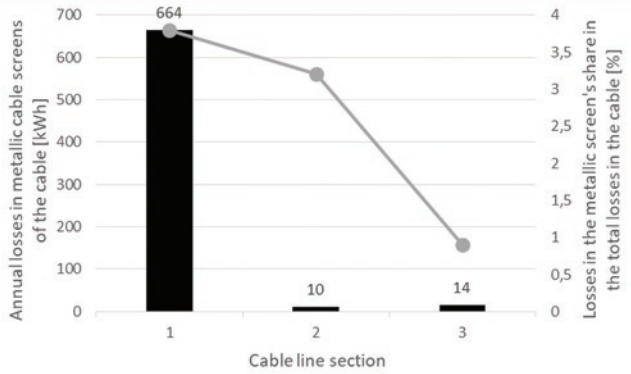


Fig. 4. Energy losses in the metallic screens of analysed cable line's sections

The similar simulations were performed for the second tested cable line. The maximum annual power output of this cable line was significantly lower and reached only 0,8 MV with the same cross-section of the cables in the analysed sections as for the first cable line analysed. The second section of the considered line was loaded at the same level as the first one, but the third section was loaded at 88% of the first section load. Due to the lower load on the second tested cable line, the absolute values of losses in the metallic screens of the cable are lower and for the entire line they amount to approx. 40 kWh per year. The share of losses in metallic screens in the total losses in individual sections of cables equals 2,3% for the first section, 1,1% for the second one and 0,5% for the third section, with the average of 1,5%.

Losses in metallic screens were also analysed in case of a busbar bridge composed of three sections of three-phase cables 3x3x240/50 mm^2 operating in parallel with a length of 23 m. The laying scheme of particular single-core cables in the tested busbar bridge is shown in Fig. 5. It is worth noting that similar cable systems can be found in lines that lead power out of large local power plants [19]. The maximum load of the bridge connected to the 25 MVA transformer was 13 MVA. After taking into consideration seasonal load changes, annual losses in metallic screens, based on computer simulation, reach the level of 150 kWh, which constitute approx. 6,7% of the total losses in the considered busbar bridge.

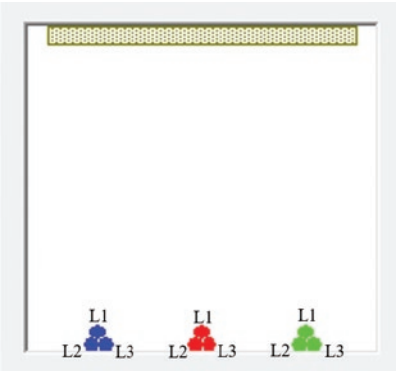


Fig. 5. Construction scheme of a busbar bridge

4. Measurement verification of simulation results

In order to verify simulation results, a series of measurements in existing MV cable lines and in a busbar bridge composed of three single-core cables were performed. The aim of the verification was to assess the discrepancy level between measured and simulated current values in metallic cable screens. Simulations were carried out using DigSILENT PowerFactory software.

Verification measurements were performed at three sections of MV cable lines in a trefoil formation and at a busbar bridge, at the substations where access to cable screens was easy. For each of the analysed cable sections, current values in core conductors and metallic screens of the cable were measured. In the measurements, the following equipment was used:

- power quality analyser Fluke 435 equipped with Rogowski coils installed on cable cores – 2% measuring accuracy,
- current clamp meters BRYMEN BM135s installed on cable screens – 5% measuring accuracy,
- in case of busbar bridges, cable core current values were recorded with four significant digit accuracy by the DMS system.

Discrepancy level was calculated with the following formula:

$$\delta I_{cs} = \frac{I_{m\_cs} - I_{s\_cs}}{I_{m\_cs}} \cdot 100\% \qquad (1)$$

where: I_{m_cs} – measured one-minute average cable screen current, I_{s_cs} – cable screen current obtained from PowerFactory simulations for the same loading conditions.

Recorded cable screen and core currents were analysed for particular load states, determined with the sampling frequency of the true RMS values recorded by the meters. Exemplary measurement results are presented in Fig. 6.

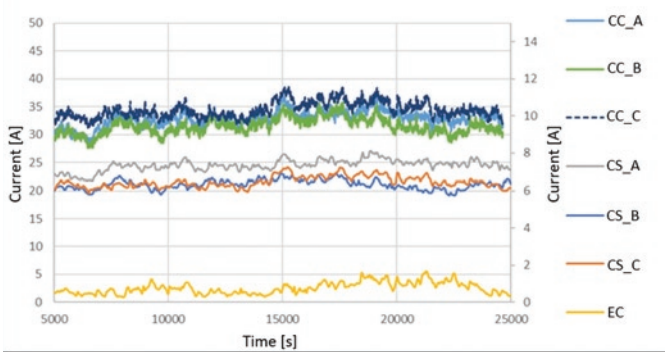


Fig. 6. Cable currents obtained from measurements (CC – cable core, CS – cable screen, EC – earth electrode current)

Significant discrepancy level between measurement and simulation results is observed for the analysed cables – measured cable screen current values are approximately +20 to +35% higher than simulation ones.. Taking into account the measurement verification of computer simulations of cable screen currents and measurement accuracy of the meters, in further considerations regarding the cost-effectiveness of the measures applied to reduce the losses in metallic screens, it is assumed that the currents flowing in the metallic cable screens are 25% higher than calculated based on computer simulations. The increased values of losses compared to the simulated ones summarized for the cable lines considered in chapter 2, presented in Fig. 7, are obtained in result of the assumption that real cable screen currents are higher by 25% than the simulated ones (for the cable line section analysed in Fig. 3).

One of the reasons for discrepancies could be harmonics content in the cable core current, transferring into cable screen current decomposition presented in Fig. 8 and not included in the PowerFactory simulation. Additional reason for discrepancies between simulation and measurement results may be the current flowing through screens, resulting from potential differences between substations' earthing installations connected by the analysed cable. The current flowing through the earth electrode, presented in Fig. 6, is the sum of above-mentioned screen currents and contains the harmonic currents as well.

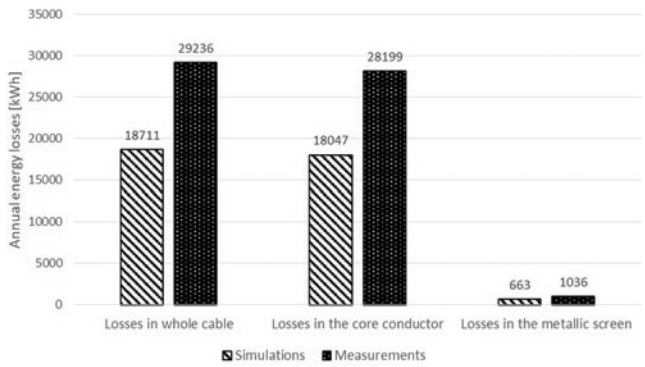


Fig. 7. Comparison of summed cable screen annual losses in investigated cable sections obtained from simulation and measurement results

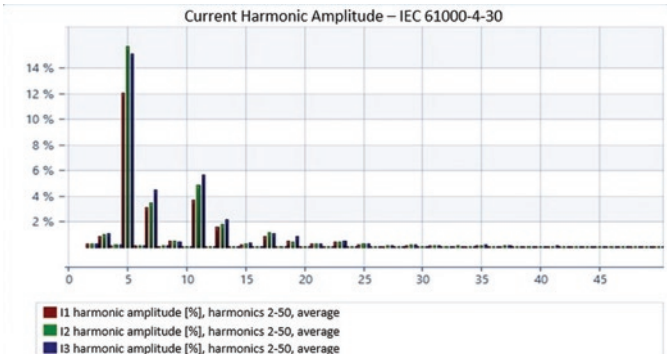


Fig. 8. Average one-month harmonic decomposition of cable screen currents in one of the analysed cables

5. Cross section selection and modifications of earthing method of metallic cable screens leading to the reduction of active power losses

The losses in metallic screens of cable line sections can be reduced by the following methods:

- cross bonding,
- one-sided earthing,
- the use of a lower cross-section compared to the standard one.

The first of the mentioned methods, when applying two cross bonding connections in one third and two thirds of the cable length, ensures reduction of losses practically to the negligible values at non-distorted currents. However, it requires the use of special cross bonding joints or the construction of cable boxes for cross bonding the screens, which is expensive. In the case of cable cross bonding the network operator's maintenance service is also expecting difficulties in conducting periodically performed cable line diagnostics.

One sided earthing of cable screens reduces losses, but requires analysis and confirmation of the admissibility of possible overvoltages of cable screen insulation at the unearthed ends of the cable screens in the case of short-circuit currents flowing through the cable cores. In addition, earthed metallic screens are used to limit the earth current flowing through the earthing installation of MV/LV substations powered by cable lines, which contributes to the limiting of possible dangerous touch voltages at these substations. Thus, in the case of cable lines operating with some of the screens earthed at one side, it is necessary to verify the values of possible touch voltages present during earth faults on these objects are within the permissible limits.

A smaller cross-section of cable screens, which is a technical solution to be applied at the stage of new cable line construction, leads to significant savings of investment expenditures. In addition, the higher resistance value of the metallic screen reduces the induced current in the both-end earthed screens, which leads to a reduction in losses. The screens with reduced cross-section may be exposed to inadmissible temperature rise during the double-phase short-circuits through the ground in the section outputting power from HV/MV substations [26]. However, to prevent this harmful thermal effect, it is possible to install instantaneous overcurrent protection for this cable section.

The use of one of the above-mentioned ways to reduce energy losses in cable screens should be economically efficient. The level of losses to which the savings can be calculated is the value of losses occurring in the cable with standard both ends earthing of screens. The proposed modification is reasonable if savings achieved thanks to the applied measure of the screen operation, exceed the investment expenditure for the cable operation modernisation during the lifetime of the cable, which can be expressed by the following relationship:

$$\Delta E_{Lmd} - V_i > 0 \tag{2}$$

where: ΔE_{Lmd} – discounted savings achieved due to modification of the cable screens operation for the considered time of cable line exploitation, V_i – value of expenditures to implement the chosen measure of the cable screens operation.

The process of decision making concerning the choice of a specific method of reducing losses in cable screens should be justified by simplified economic analyses using relatively easily available data characterising the analysed cable line. In order to evaluate the possible savings on losses in the cable screens, computational simulations of losses in these screens were performed in the function of cable loading for the considered loss reduction methods, such as cross bonding, single-sided screen earthing and the use of reduced screen cross-section. The values of cable screen current calculated in computer simulations were corrected according to the results of measurement verification of simulation results, which revealed an average increase of 25% in current values flowing in the screens operating within the distribution network, in relation to the values obtained in computational simulations. This can be translated into the following increase in losses in real metallic screens (3):

$$\Delta P_{re} = I_{re}^2 \cdot R = (1{,}25 \cdot I_{sim})^2 \cdot R = 1{,}5625 \cdot I_{sim}^2 \cdot R = 1{,}5625 \cdot \Delta P_{sim} \tag{3}$$

where: ΔP_{sim} – value of power losses obtained from simulation, I_{sim} - current in the cable screens obtained from simulation, R - cable resistance, I_{re} – measured current in the cable screens, ΔP_{re} - actual losses in the cable screens.

When cable screen cross bonding is applied, the losses in screens are negligible. For other cases analysed, the obtained results showing values of losses in the cable screens in relation to the losses in the core conductors for selected cable and screen cross sections are presented in Fig. 9. The visible shaded range of losses values for individual cable types results from different lengths of analysed cable sections in the range from 0,1 to 1 km. It was found, as the overall result of simulations conducted, that the share of losses in cable screens in the losses in the core conductors of the analysed cable systems is constant in the range from 5 to 100% of the rated current load, which is illustrated in Fig 9.

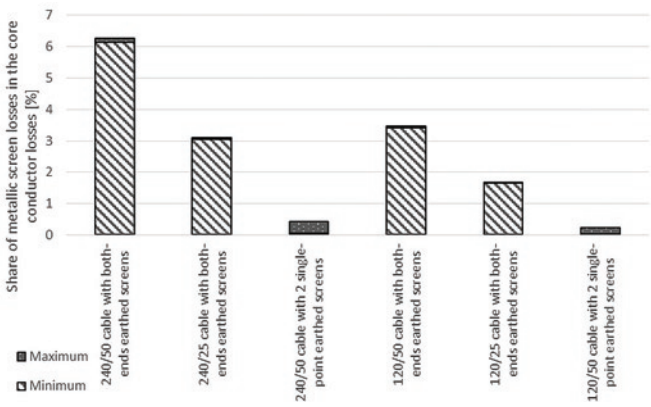


Fig. 9. The share of power losses in cable screens in the losses in core conductors analysed for cases of various cable screen operation and cross section

In the case of short sections of cable lines operating in substation cable bus bar bridges, cross bonding is not applicable, however, single point earthing of cable screens is rational since cable screens operation mode does not influence the electrical shock hazards at the HV/MV after cable screen proper insulation at the unearthed ends. The shares of

power losses in cable screens in core conductor losses in busbar bridges for various cable screen connection configurations obtained on the basis of the measurement corrected simulation of the considered busbar bridges operation at various loads are shown in Fig. 10.

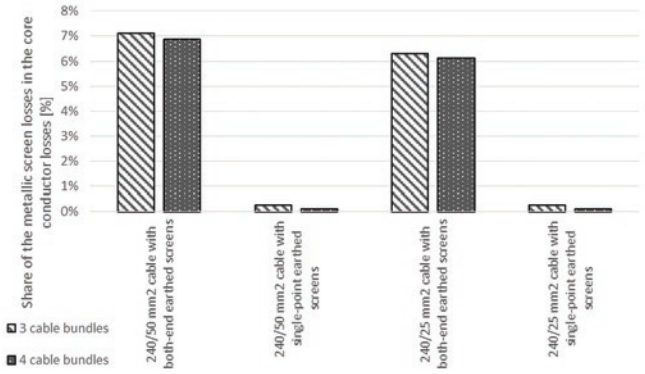


Fig. 10. The share of power losses in cable screens in core conductor losses in cable busbar bridges for various cable screen connection configurations and cable busbar bridges composed of 3 or 4 cable bundles

The presented results can be used to estimate the annual savings resulting from the change in the method of cable screens operation in accordance with the following procedure.

A. Determining the maximum losses P_{max} in the cable core conductors (4):

$$P_{\max} = 3 \cdot I_{\max}^2 \cdot \frac{l}{\gamma \cdot s} \tag{4}$$

where: I_{max} - current in the core conductor of the analysed cable section, l - length of the section [m], γ - conductivity of the cable core [m/(Ωmm^2)], s - cross-section [mm^2].

The load data of the analysed sections can be obtained in the case of power line bays in HV/MV substations from the registered values by the distribution network management system (DMS); for cable sections situated down the power line, when measured load values are not available, the approximate load can be calculated by multiplying the power leaded out from a given line bay of HV/MV substation, registered by DMS, by the share of the sum of rated power of MV/LV transformers supplied from a given section of cable line in the summed rated power of all transformers supplied from the power line supplied by the given bay in the analysed substation.

B. Determining the annual savings in energy losses $\Delta E_{l(BEB\text{-}m)i}$ in cable screens during a year as a result of the change in the way cable screen operation from both end earthing (*BEB*) system to the new modernised way (*m*):

$$\Delta E_{l(BEB-m)i} = (\Delta P_{lBEB\%} - \Delta P_{lm\%}) \cdot P_{\max} \cdot \tau_{\max} \cdot \left(1 + \frac{\Delta P_a}{100}\right)^i \tag{5}$$

where: $\Delta P_{lBEB\%}$; $\Delta P_{lm\%}$ - values of power losses in cable screens as percentage of the total cable power losses for the considered ways of cable screen operation presented as samples in Fig. 9 and 10; P_{max} - losses in the core conductors of the cable line according to the equation (4), τ_{max} - annual duration of maximum losses in the analysed cable section, ΔP_a - annual increase of losses due to the average increase in annual load of the section in [%].

In order to determine the power flows in individual sections of cable lines P_{max}, one may use the method of load distribution based on rated power of supplied transformers, load distribution based on monthly readings of cable section energy flow, analysis of load profiles or use of data recorded by smart meters (AMI, Advanced Metering Infrastructure) [1]. It is widely assumed that the use of data from the AMI system allows to achieve the most accurate results [25, 47], however due to the fact that the AMI system does not achieve full functionality, the data based on standard load profiles [22] are often used in the analyses.

The annual durations of maximum losses τ_{max} for segments of the distribution network in Poland can be found in the bibliography and according to [36], for bus bar bridges and cable sections leading out power from HV/MV substation bays, their value range is 1248 - 4449 h with an average value of 2525 h and a standard deviation value of 510 h. Selection of the right duration time of maximum losses depend on the types of loads supplied, greater for industrial loads and smaller for residential ones. For sections situated down the line, loaded with MV/LV substations, the range of considered duration times lies in the value range from 788 to 2444 h with an average value of 1662 h and standard deviation of 801 h. For power lines supplying individual MV/LV substations deep in the distribution network, when the results of load measurement are not available, it will be rational to adopt the values from the given range of values in proportion to the summed rated power of MV/LV substations supplied from a given section in relation to the total summed rated power of substation transformers supplied from the analysed line taking into account the types of loads supplied as well.

C. Discounted value of losses for cable line lifetime ΔE_{lmd}

In order to determine the value of discounted losses saved ΔE_{lmd}, the annual savings resulting from modernization, after discounting their values to the initial year level, should be summed up for the cable lifetime considered using the following relationship:

$$\Delta E_{Lmd} = \sum_{i=1}^{25} \left[\Delta E_{l(BEB-m)i} \cdot C_{ee0} \cdot \left(1 + \frac{\Delta C_a}{100}\right)^i \cdot \left(1 + \frac{R_d}{100}\right)^{-i} \right] \tag{6}$$

where: $\Delta E_{l(BEB\text{-}m)i}$ - annual savings in energy losses given by (5), C_{ee0}, ΔC_a - market price of electric energy in the year zero [PLN/kWh] and its expected annual growth in [%], R_d - annual discounting rate used in [%] to determine the value of losses in initial year.

The discount period was assumed to be 25 years, i.e. slightly above the accounting depreciation period of the cable lines in Poland amounting to 22,5 years [45].

Investment expenditure value V_i for the modernization of the cable screens operation ways, presented in inequality (2), depend on the solution applied. The cheapest solution is to isolate the cable screens at one end, which is related to the costs of a service brigade operation carrying out such insulation and the rather low equipment costs. The use of cross bonding is definitely more expensive, because it requires the purchase and installation of cross bonding joints or cable boxes for the implementation of this solution. The use of a metallic screen with a smaller cross-section compared to that considered as standard solution leads to significant savings in investment expenditure, the more important the longer the analysed section of the cable is and the limitation of the loss level as well. However, this can only be implemented at a stage of the cable line construction because the replacement for a cable with a smaller cross-section of metallic screen would not be economically efficient, due to the high investment costs close to the initially spend value for the construction of the cable line.

Numerous analyses of the cost-effectiveness of limiting losses by applying the above-mentioned modernization ways of cable screen operation have been carried out. The cost of losses were discounted to the sample initial investment year i.e. 2017 using two discount rates: a low value of 2,85% used in calculating the value of public assistance

applied for example in the distribution of due payments in instalments and a higher having value of 5,633% being the return rate of capital cost obtained by distribution system operators in Poland. An increase in the annual cable line load was assumed to be 0,5% and an annual increase in electricity prices to cover losses at the level of 2,5% annually in relation to the 2017 price. The results of economic analyses performed lead to the following conclusions:

- lack of profitability of cable screens cross bonding for currently observed cable power line loads in the distribution network supplying residential and commercial customers due to the low value of discounted savings on energy losses reduction which for the analysed line segments amount up to approximate sum of PLN 3 000 over the period of 25 years,
- the cost-effectiveness of applying the one single-point earthing for two out of 3 cable screens of cable power line loaded with maximum power greater than 2 MVA for residential consumers,
- lack of profitability of applying single side earthing of cable screens in existing cable bus bar bridges due to the low value of discounted losses saved in these bridges, which results from their small length, despite their significant load,
- profitability of using 25 mm^2 cable screens in place of 50 mm^2 used as standard so far in new cable power lines and considering the profitability of applying single side earthing of 2 out of 3 cable screens in such cables depending on the results of simplified economic analysis,
- profitability of using 25 mm^2 cable screens in cable bus bar bridges and the application of single point earthing of all of their metallic cable screens in the case of new investments.

6. Risks resulting from cable screens operation with single point earthing

Single point earthing of some or all cable screens is the most interesting way of reducing cable screen losses. Unfortunately, some risks resulting from the single point earthing of cable screens are identified and presented below.

6.1. Risk of electrical electrocution in MV/LV substations

Single point earthing of cable screens has impact on earthing currents flow and contact voltage, which occur in MV/LV substations under phase to earth fault conditions and in LV circuits supplied from these substations [41, 42]. Earth fault current I_{k1} is divided into two paths – current flowing through earthing system of MV substation I_e and return current flowing through cable screens I_{cs}:

$$I_{k1} = I_e + I_{cs} \tag{7}$$

Only I_e component causes increase of contact voltage in substations. Decrease of the earthing current I_e due to the current flowing partially through the cable screen is presented by the value of reduction factor r, causing reduction of the earthing current I_e:

$$I_e = r \cdot I_{k1} \tag{8}$$

The reduction factor of a cable line consisting of three single core cables is usually calculated with the formula:

$$r = 1 - \frac{Z_M}{Z_{CS}} = \frac{Z_{CS} - Z_M}{Z_{CS}} \tag{9}$$

where: Z_M – mutual coupling impedance between cable core with phase to earth fault current and cable screens; Z_{CS} – self impedance of cable screen.

Formula (9) is fully valid in case of neglecting the earthing resistance of three cable screens. Taking into account the resistance of earthing systems [5], the resulting earth fault current flow is presented in Fig. 11, for the case of earth fault at the end of the cable line at MV/LV substation, supplied directly from HV/MV substation.

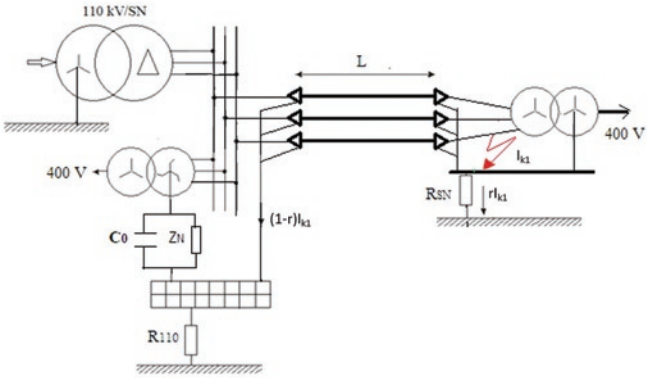


Fig. 11. Phase to earth current flow during the earth fault at MV/LV substation, including cable screens of the cable, as well as resistance of earthing systems of both substations; Z_N – impedance of MV neutral point earthing (resistor or Petersen's coil), C_0 – earth fault network capacitance, I_{k1} – earth fault current, R_{110} – earthing resistance of the HV substation, R_{MV} – earthing resistance of the MV substation, L – length of the considered cable line [km]

Voltage difference between earthing systems of HV/MV substation and MV/LV substation can be described by the following formula:

$$(1-r)I_{k1}Z_{CSu} - I_{k1}Z_{Mu} \cdot L = rI_{k1}(R_{110} + R_{MV}) \tag{10}$$

where: Z_{Mu} and Z_{CSu} - unit impedances [Ω /km], r – reduction factor, R_{110} – *earthing resistance of the HV substation*, R_{MV} – *earthing resistance of the MV substation*, L - cable length [km].

After conversion of (10), the formula describing the earthing current is obtained:

$$I_e = I_{k1} \cdot \frac{(Z_{CSu} - Z_{Mu}) \cdot L}{(R_{110} + R_{MV}) + Z_{CSu} \cdot L} \tag{11}$$

Due to the fact that most often $R_{110} << R_{MV}$, in the calculations only the earthing resistance of the supplied substation R_{SN} is taken into account.

It can easily be noticed that formula can be transformed to the following version:

$$I_e = r \cdot I_{k1} \cdot \frac{Z_{CSu} \cdot L}{R_{MV} + Z_{CSu} \cdot L} \tag{12}$$

or

$$I_e = r_{re} \cdot I_{k1} \tag{13}$$

where:

$$r_{re} = r \cdot K_{cor} \tag{14}$$

and the correction factor K_{cor}:

$$K_{cor} = \frac{Z_{CSu} \cdot L}{R_{MV} + Z_{CSu} \cdot L} \tag{15}$$

Table 1. Change of reduction factor values resulting from single point earthing of cable screens as multiplicity of reduction factor for both-end earthed screens under assumption that impedances Z_{CS} and the sum of earthing resistances (R_{110}+R_{MV}) are comparable.

No.	Growth rate of r value according to measurement results	Growth rate of r value according to analytical calculations	Growth rate of r value according to PF simulations results	Remarks
1	1,20	1,28	1,20	for one single point cable screen earthed in substation HV/MV
2	1,70	1,79	1,75	for two single point cable screens earthed in substation HV/MV

Correction factor K_{cor} presented in the formula (15) is valid for cables in which three cable screens are earthed. After single point earthing of one or two cable screens in MV/LV substation, value of that factor is changed because impedance Z_{CS} is changed. Expected variation range of the reduction factor values for such cases is presented in the table 1 for three different ways of determining the value of correction factor:

- measurement results on real cable line supplied with 230/400 V,
- simplified analytical calculations, which include cable formation and cable parameters influence,
- computer simulations using PowerFactory software.

When assessing the earthing electrode currents flow in MV substations supplied by cable lines, the impact of the substation earthing electrode resistance should also be taken into account in the calculation of the reduction factor. For the assumed sum of earthing resistances (R_{110}+R_{MV}) in range 0 – 5 Ω, the correction factor of the reduction factor values change can be illustrated by the curve depicted in Fig. 12. In order to determine the resistance of the earthing electrodes of the actual MV/LV substation, measurements should be taken in accordance with [13, 14].

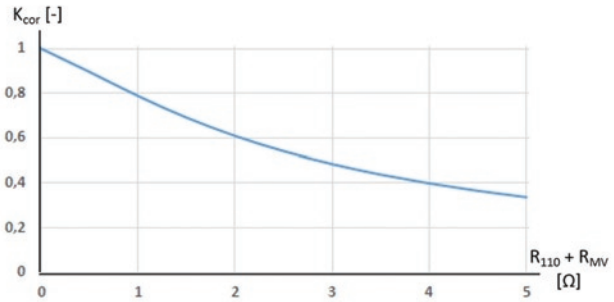


Fig. 12. Changes of reduction factor correction coefficient in function of the sum of earthing resistances of cable screens

Simulation and measurement results on cable line supplied with the reduced voltage showed that reduction factor increases when cable screens are unearthed. For the analysed case, the value of reduction factor equals 0,5 for both-end earthed cable and 0,6 when one cable screen is single point earthed or 0,85 when two cable screens are single point earthed. Permissibility of cable screen single point earthing operation should be therefore verified based on acceptable touch voltage level, which depend also on earth electrodes' resistance values at substations connected by the cable screens. Data for such analysis can be obtained by simulating earth fault current flow for the actual neutral point impedance, the mode of cable screens operation and earthing resistance of cable screens involved.

In case of cable bus bar bridges earthing at both ends or single point earthing does not have significant impact on a change of current flowing through earthing system of the HV/MV substation and therefore does not have a significant impact on risk of electric shock during earth faults at this substation.

6.2. Overvoltages at the end of cable line with single point earthed cable screens

Single point earthing of cable screens does not result in overvoltages under normal loading conditions. Hazard is created during phase to earth faults and phase to phase faults in MV cables, where, as it appears in simulation, peak value of transient overvoltage at the unearthed end of cable screen can be in range of over a dozen kV, which could be dangerous to cable sheath insulation. Not only does the value of transient overvoltages have influence on a risk of cable sheath damage, but also a frequency of fault occurrence in cable line. The most frequent faults can be observed in MV lines composed of several cable and overhead line sections, because of big intensity of faults in overhead lines. As a result, negative cumulative aging effect of cable screen insulation can be expected [31, 39]. In case of single point earthing of two cable screens, it could be beneficial, in order to reduce transient overvoltages, to bond those two unearthed screen ends, which should decrease the change in surge impedance of the screen and possibly result in the reduced overvoltages, but at the same time, it causes losses' increase in the connected screens.

According to simulation results, both side earthed screens reduce transient and steady-state overvoltages at the remaining single point earthed screens in comparison with the overvoltages in case of single-point earthing of all of cable screens. It can be supposed that both side earthed screen act in a similar way as ECC cable laid in proximity of HV cable lines applied there for reducing overvoltages [51, 52].

Application of cable screen single point earthing in MV cable lines requires further investigations in terms of transient overvoltages under phase to earth fault conditions and analysis of efficiency of overvoltage reduction methods, as well as the impact of overvoltages on cable screen insulation.

Simulation results indicate that overvoltages in cable bus bar bridges under phase to earth fault taking place in cable feeders are characterized by relatively low peak value, however the occurrence of overvoltages is more frequent, i.e. they occur during every short circuit in a network supplied by a given HV/MV substation. Theoretically cable bus bars could be safely unearthed because peak transient overvoltages values are below 2 kV even during phase to phase short-circuits through the earth and therefore is below 5 kV used under their DC voltage acceptance tests.

7. Conclusions

The growing share of cable lines, composed of single-core cables with metallic screens, in the overall length of medium voltage power distribution network, encourages the analysis of feasibility of possible changes of their standard operation with the aim to find the optimal solutions. In particular the choice of operation mode preventing excessive losses in the cable metallic screens seems to be important. In the presented analysis the value of losses in metallic screen of cable lines and bus bar power bridges with both side earthing of their screens was investigated in function of the loading carried for residential and commercial load profile. The computer simulations of cable line sections operation were performed and their results were verified concerning the energy losses in cable screens by measurements in real objects.

This verification indicated a 25% underestimation of the currents induced in the cable metallic screens obtained in computer simulations. One of the reasons for higher losses may be the content of harmonic distortions of currents in conductors.

The conducted simulations, corrected in result of measurement verification, can be the basis for the assessment of financial losses resulting from the both side earthing of cable metallic screens during the cable lifetime operation. The level of economic losses assessed for such standard operation of cable screens may lead to a modification of the way these screens are operated by applying remedial measures such as cross bonding, single side earthing or the use of a smaller cross-section of cable screens. The simplified methodology for assessing the economic value of reduction of losses possible to be achieved is proposed in this article, based on relationships enabling to calculate the savings on losses in 25 years of cable operation. The value of losses is obtained by discounting annual values to the initial year of analysis assuming a specific annual increase in energy prices and of cable loading rate over the analysed period. The determined savings constitute the basis for making a decision about implementation of possible modification of cable screens operation mode or for the proper choice of cable screens cross section.

The basic conclusion from the conducted analysis is the recommendation to limit the cross-section of cable metallic screens to 25 mm^2. At the current load level of cable lines in the distribution network supplying residential and commercial facilities, the level of discounted losses does not justify the use of quite costly remedies in the form of cable screens cross bonding.

A relatively simple remedy is the operation of cable screens with single point earthing. Application of this solution is recommended for all cable screens in the case of new installations of bus bar cable bridges made of single-core cables. In the case of cable lines, in cases demonstrating the profitability of limiting losses by applying single point earthing of one or two cable screens, such a solution can be conditionally applied leaving always one earthed screen on both ends to ensure the possibility of short-circuit current reduction.

The article describes the problem of the reduction of earth fault current following the single point earthing of metallic cable screens. The results of the reduction of current flowing to the ground through earth electrodes based on analytical equations are presented, as well as computer simulations and tests of cables laid in the ground supplied with reduced voltage. Procedures for determining the distribution of short-circuit currents should take into account the resistances of cable screen conductor and the resistances of earthing installations in the substations to which the cable screen conductors are bonded. In urban areas where earthing systems can be considered as forming the integrated earthing installations, the limited reduction of short circuit currents should not be an obstacle in applying the analysed remedy to reduce losses.

Computer simulations of overvoltages of the metallic screen outer cable insulation in cases of single point earthing of metallic screens that may occur during short-circuits are performed. Methods that can limit the values of such overvoltages are proposed for further research purposes. With two single side earthed screens, it is proposed to short-circuit the single point earthed screens at the unearthed end, resulting in the unchanged screen impedance for overvoltage waves, but for the price of increased losses similar to the case of one single point earthed screen. An alternative solution may be leaving two screens both side earthed, which can result in similar overvoltage's reduction as observed in high voltage cable equipped with ECC wire.

References

1. Arritt R, Dugan R. Comparing Load Estimation Methods for Distribution System Analysis. 22nd International Conference and Exhibition on Electricity Distribution. CIRED 2013, https://doi.org/10.1049/cp.2013.0869.
2. Barrett J S, Anders G J. Circulating current and hysteresis losses in screens, sheaths and armour of electric power cables - mathematical models and comparison with IEC Standard 287. IEE Proceedings - Science, Measurement and Technology 1997; 144: 101-110, https://doi.org/10.1049/ip-smt:19971162.
3. Becker J, Musique F. Economical 36 kV Cable System for the Belgian Network. 16th International Conference and Exhibition on Electricity Distribution. CIRED 2001, https://doi.org/10.1049/cp:20010673.
4. CEER Benchmarking Report 6.1 on the Continuity of Electricity and Gas Supply. Council of European Energy Regulators asbl 2018.
5. Charlton T E, Hocaoglu M H, Karacasu O, Marican A M A. Impact of cable sheath sizing, material and connections upon the safety of electrical power installations. 21st International Conference on Electricity Distribution. CIRED 2011.
6. Czapp S, Dobrzyński K, Klucznik J, Lubośny Z. Analiza napięć indukowanych w żyłach powrotnych kabli wysokiego napięcia dla ich wybranych konfiguracji. XVIII Międzynarodowa Konferencja Aktualne Problemy w Elektroenergetyce APE 2017.
7. Deschamps P, Toravel Y. Reduction of Technical and Non-Technical Losses in Distribution Networks. CIRED Overview, final report 2017.
8. DIgSILENT Power Factory 2018 Technical Reference Documentation. Cable System, ElmCabsys, TypCabsys.
9. DIgSILENT Power Factory 2018 User Manual.
10. Dong X, Yang Y, Zhou C, Hepburn D M. Online Monitoring and Diagnosis of HV Cable Faults by Sheath System Currents. IEEE Transactions on Power Delivery 2017; 32: 2281-2290, https://doi.org/10.1109/TPWRD.2017.2665818.
11. Duda D, Szadkowski M. Ochrona przeciwprzepięciowa osłon kabli WN w różnych układach połączeń żył powrotnych. Przegląd Elektrotechniczny 2014; 10: 37-40.
12. Elektroenergetyczne linie kablowe średniego napięcia. Standard w sieci dystrybucyjnej Enea Operator Sp. z o.o. 2017.
13. European Standard PN-EN 50341-1:2013-03 Overhead electrical lines exceeding AC 1 kV - Part 1: General requirements - Common specifications.
14. European Standard PN-EN 50522:2011 Earthing of power installations exceeding 1 kV a.c.
15. European Standard PN-EN 61936-1:2011 Power installations exceeding 1 kV a.c. - Part 1: Common rules.
16. European Standard PN-HD 60364-4-442:2012 Electrical Installations of Buildings -Protection for Safety - Protection of Low-Voltage Installations Against Temporary Overvoltages and Faults Between High-Voltage Systems and Earth.
17. Gołaś A, Ciesielka W, Szopa K, Zydroń P, Bąchorek W, Benesz M, Kot A, Moskwa S. Analysis of the possibilities to improve the reliability of a 15 kV overhead line exposed to catastrophic icing in Poland. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(2): 282-288, https://doi.org/10.17531/ein.2019.2.13.

18. Gouda O E, Farag A A. Factors affecting the sheath losses in single-core underground power cables with two-points bonding method. International Journal of Electrical and Computer Engineering (IJECE) 2012; 2: 7-16, https://doi.org/10.11591/ijece.v2i1.115.
19. Gouramanis K V, Kaloudas Ch G, Papadopoulos T A, Papagiannis G K, Stasinos K. Sheath Voltage Calculations in Long Medium Voltage Power Cables. IEEE Trondheim PowerTech 2011, https://doi.org/10.1109/PTC.2011.6019234.
20. Heiss A, Balzer G, Schmitt O, Richter B. Surge arresters for cable sheath preventing power losses in MV networks. 16th International Conference and Exhibition on Electricity Distribution. CIRED 2001, https://doi.org/10.1049/cp:20010706.
21. IEEE Guide for the Application of Sheath-Bonding Methods for Single-Conductor Cables and the Calculation of Induced Voltages and Currents in Cable Sheaths. IEEE Standard 575 1988.
22. Instrukcja Ruchu i Eksploatacji Sieci Dystrybucyjnej Enea Operator Sp. z o.o.: Standardowe profile zużycia energii na rok 2017. http://www.operator.enea.pl/21/instrukcje/instrukcje-iriesd-883.html, available: 24.07.2019.
23. Jensen C F, Nanayakkara O M K K, Rajapakse A D, Gudmundsdottir U S, Bak C L. Online fault location on AC cables in underground transmissions systems using sheath currents. Electric Power Systems Research 2014; 115: 74-79, https://doi.org/10.1016/j.epsr.2014.04.002.
24. Jung C K, Lee J B, Kang J W, Wang X H, Song Y H. Sheath Current Characteristic and Its Reduction on Underground Power Cable Systems. IEEE Power Engineering Society General Meeting 2005.
25. Kauppinen M, Pylvanainen J, Karjalainen J, Sihvola V, Experiences of using AMI system for DSO's business operation. CIRED - Open Access Proceedings Journal 2017; 1: 2756-2759, https://doi.org/10.1049/oap-cired.2017.0764.
26. Korab R, Siwy E. Statistical analysis of the double line-to-ground short-circuit current in MV urban network for the power cable metallic screen rating. International Conference on Probabilistic Methods Applied to Power Systems 2006, https://doi.org/10.1109/PMAPS.2006.360342.
27. Korab R, Siwy E, Żmuda K. Analiza możliwości redukcji przekroju żył powrotnych w kablach średniego napięcia w sieciach miejskich. Zeszyty Naukowe Politechniki Śląskiej. Elektryka 2004; 189: 111-120.
28. Korab R, Siwy E, Żmuda K. Podstawy racjonalnego doboru żył powrotnych (ekranów) kabli 6-20 kV różnego typu w sieciach miejskich. Elektroenergetyczne linie kablowe 2004; 148-149: 19-28.
29. Kornatka M. Analysis of the exploitation failure rate in Polish MV networks. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(3): 413-419, https://doi.org/10.17531/ein.2018.3.9.
30. Krawiec H. Przyczyny grzania się bednarki i żył powrotnych kabli 6 kV. Zeszyty Problemowe - Maszyny Elektryczne 2014; 1: 141-145.
31. Li J, Xu L, Chen X, Zhao A, Liu J, Zhao X, Deng J, Zhang G. Analysis of Statistical and Frequency Characteristics of Transient Overvoltage of Hybrid Cable-OHL Lines. China International Conference on Electricity Distribution (CICED) 2018; 2650-2654, https://doi.org/10.1109/CICED.2018.8592463.
32. Lin Y, Xu Z. Cable Sheath Loss Reduction Strategy Research Based on the Coupled Line Model. IEEE Transactions on Power Delivery 2015; 30: 2303-2311, https://doi.org/10.1109/TPWRD.2015.2414655.
33. Lin Y, Yang F, Xu Z, Weng H. Cable Sheath Loss Analysis Based on Coupled Line Model. International Conference on Power System Technology 2014.
34. Łowczowski K. Badanie wpływu ułożenia kabli na straty energii w żyle powrotnej - symulacja w programie Power Factory. Przegląd Elektrotechniczny 2016; 10: 54-57, https://doi.org/10.15199/48.2016.10.13.
35. Moore G F. Electric Cables Handbook, Third Edition. Oxford: Blackwell Science Ltd, 1997.
36. Niewiedział R, Niewiedział E, Wyznaczanie czasu trwania strat maksymalnych w sieciach elektroenergetycznych modelami obliczeniowymi. http://www.sep.krakow.pl/nbiuletyn/nr57ar2.pdf, available 5.07.2019.
37. Novak B, Koller L, Berta I. Loss reduction in cable sheathing. Renewable Energy and Power Quality Journal 2010; 1: 293-297, https://doi.org/10.24084/repqj08.311.
38. Novak B, Koller L. Current distribution and losses of grouped underground cables. IEEE Transactions on Power Delivery 2011; 26: 1515-1521, https://doi.org/10.1109/TPWRD.2011.2104980.
39. Orsagova J, Toman P. Transient overvoltages on distribution underground cable inserted in overhead line. 20th International Conference on Electricity Distribution. CIRED 2009, https://doi.org/10.1049/cp.2009.0883.
40. Polityka Energetyczna Polski do 2040 roku, projekt. Warszawa: Ministerstwo Energii, 2018.
41. Pons E, Colella P, Napoli R, Tommasini R. Impact of MV Ground Fault Current Distribution on Global Earthing Systems. IEEE Transactions on Industry Applications 2015; 51: 4961-4968, https://doi.org/10.1109/TIA.2015.2417836.
42. Popović L M. Ground Fault Current Distribution When a Ground Fault Occurs in HV Substations Located in an Urban Area. Progress In Electromagnetics Research B 2014; 59: 167-179, https://doi.org/10.2528/PIERB14022004.
43. Rakowska A. Service Experiences for MV Cable Network - Optimistic or Pessimistic State of The Art. Jicable Conference: The leading forum about Insulated Power Cables 2007.
44. Riba Ruiz J R, Garcia A, Alabern Morera X. Circulating sheath currents in flat formation underground power lines. 2007, https://doi.org/10.24084/repqj05.217.
45. Rozporządzenie Rady Ministrów z dnia 10.12.2010 r. w sprawie Klasyfikacji Środków Trwałych (KŚT).
46. Sakalkale S, Kanakgiri K. Study of Underground Power Cable Considering Sheath Circulating Current. Proceedings of Fourth IRF International Conference 2014; 7-10.
47. Sapienza G, Noce C, Valvo G. Network Technical Losses Precise Evaluation Using Distribution Management System and Accurate Network Data. 23rd International Conference on Electricity Distribution. CIRED 2015.
48. Sobral A, Moura A, Carvalho M. Technical Implementation of Cross Bonding on Underground High Voltage Lines Projects. 21st International Conference on Electricity Distribution. CIRED 2011.
49. Sullivan R G, Abplanalp J M, Lahti S, Beckman K J, Cantwell B L, Richmond P. Electric Transmission Visibility and Visual Contrast Threshold Distances in Western Landscapes. National Association of Environmental Professionals Annual Conference 2014; 1-46.
50. Tarko R, Benesz M, Nowak W, Szpyra W. Statystyczna analiza zakłóceń zwarciowych dla określenia przekroju żył powrotnych kabli średnich napięć. Przegląd Elektrotechniczny 2016; 7: 186-189, https://doi.org/10.15199/48.2016.07.40.
51. Todorovski M, Ackovski R. Equivalent Circuit of Single-Core Cable Lines Suitable for Grounding System Analysis Under Line-to-Ground

Faults. IEEE Transactions on Power Delivery 2014; 29: 751-759, https://doi.org/10.1109/TPWRD.2013.2277887.
52. Żmuda K. Elektroenergetyczne układy przesyłowe i rozdzielcze. Wybrane zagadnienia z przykładami. Gliwice: Wydawnictwo Politechniki Śląskiej, 2016.

Jerzy ANDRUSZKIEWICZ
Józef LORENC
Krzysztof ŁOWCZOWSKI
Agnieszka WEYCHAN
Institute of Electric Power Engineering
Poznan University of Technology
ul. Piotrowo 3a, 60-965 Poznań, Poland

Józef ZAWODNIAK
Enea Operator sp. z o.o.
ul. Strzeszyńska 58, 60-479 Poznań, Polska

E-mails: jerzy.andruszkiewicz@put.poznan.pl, jozef.lorenc@put.poznan.pl, krzysztof.lowczowski@put.poznan.pl, agnieszka.weychan@put.poznan.pl, jozef.zawodniak@enea.pl

Article citation info:
ZHANG C, ZHANG Y. Common cause and load-sharing failures-based reliability analysis for parallel systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 26–34, http://dx.doi.org/10.17531/ein.2020.1.4.

Chao ZHANG
Yunan ZHANG

COMMON CAUSE AND LOAD-SHARING FAILURES-BASED RELIABILITY ANALYSIS FOR PARALLEL SYSTEMS

ANALIZA NIEZAWODNOŚCI SYSTEMÓW RÓWNOLEGŁYCH W SYTUACJI JEDNOCZEŚNIE WYSTĘPUJĄCYCH USZKODZEŃ WYWOŁANYCH WSPÓLNĄ PRZYCZYNĄ ORAZ USZKODZEŃ ELEMENTÓW DZIELĄCYCH OBCIĄŻENIE

For parallel system reliability, the mean time to failure of parallel system under common cause failure (load-sharing failure) is shorter than that of the system without common cause failure (load-sharing failure). The traditional calculation approaches of mean time to failure of parallel systems do not consider the possible effect of common cause and load-sharing failure. However, it may result in the poor accuracy of mean time to failure of parallel system and pose a threat to system reliability. This paper not only considers the effect of common cause failure with stress strength, but also investigates the joint effect of the load-sharing and common cause failures. Besides, the joint failure model of three-dependent-component parallel system are established, and the corresponding properties are analyzed. Finally, a numerical example is used to illustrate the proposed method.

***Keywords**: system reliability, mean time to failure, common cause failure, load-sharing failure, parallel system.*

Gdy mowa o niezawodności systemu równoległego, średni czas do uszkodzenia, w przypadku uszkodzenia wywołanego wspólną przyczyną (lub uszkodzenia elementów dzielących obciążenie) jest krótszy niż dla systemu, w którym nie występują tego typu uszkodzenia. Tradycyjne metody obliczania średniego czasu do uszkodzenia systemów równoległych nie uwzględniają potencjalnego wpływu uszkodzeń wywołanych wspólną przyczyną oraz uszkodzeń komponentów dzielących obciążenie. Może to skutkować małą dokładnością tak obliczanego średniego czasu do uszkodzenia systemu równoległego i stanowić zagrożenie dla jego niezawodności. W prezentowanej pracy rozważano nie tylko wpływ uszkodzenia wywołanego wspólną przyczyną dla modelu typu wytrzymałość-obciążenie, ale również wpływ jednocześnie występujących uszkodzeń wywołanych wspólną przyczyną i uszkodzeń elementów dzielących obciążenie. Poza tym opracowano model, w którym omawiane dwa typy uszkodzeń występują jednocześnie w systemie równoległym składającym się z trzech zależnych elementów oraz przeanalizowano właściwości takiego systemu. W artykule przedstawiono przykład numeryczny, który ilustruje zastosowanie proponowanej metody.

***Słowa kluczowe**: niezawodność systemu, średni czas do uszkodzenia, uszkodzenie wywołane wspólną przyczyną, uszkodzenie elementów dzielących obciążenie, system równoległy.*

1. Introduction

A. Background

Today's systems are becoming more complex and more sophisticated, and the problems of system reliability are drawing an increasing attention. Common cause failures are critical risk contributors in complex technological systems as they challenge multiple redundant systems simultaneously. Common cause failures can contribute significantly to the overall system unreliability [9]. Therefore, it is important to incorporate common cause failure into the system reliability analysis. Alizadeh et al. [1] introduced the impact of common cause failure on the system reliability using Markov analysis technique. Zuo et al. [23] analyzed the system failure suffering common cause failure. Fan et al. [2] developed a new model for common cause failures considering components degradation based on mathematical framework of Stochastic Hybrid Systems. Levitin [5] adapted the universal generating function method of multistate system reliability analysis to incorporate common-cause failures. Pourali [7] presented presented the importance of considering common cause failure in reliability, availability, and maintainability analysis for industrial and commercial mission-critical facilities and high-reliability organizations. Vaurio [11] incorporated common-cause failures into system analysis by an implicit method and discussed the possible limitations and extensions. Wang et al. [13] incorporated effects of probabilistic common cause failures into system reliability analysis. Wang et al. [14] proposed an explicit method and an implicit method to analyze the reliability of systems. Xiao and Gao [15] proposed efficient simulation methods to assess the system reliability with input uncertainty. Xiao et al. [16] presented a data simulation approach to estimating the system failure probability in the presence of stochastic constraints. Yuan [17] extended the pivotal decomposition method for system availability and failure frequency from the case where components are statistically independent to that where components are also subject to common-cause failures.

Load-sharing is always an essential nature in parallel system. Huang et al. [3] presented a general closed-form expression for lifetime reliability of load-sharing redundant systems. Liu [6] developed a model to calculate the reliability of a load-sharing system which is composed of non-identical components each having an arbitrary failure time distribution. Paula et al. [8] analyzed the optimization in redundant system considering load sharing. Jiang et al. [4] formulated two load optimization models to identify the optimal loading strategy. Sutar et al. [10] modeled the load sharing phenomenon in a k-out-of-

m system through the accelerated failure time model. Wang et al. [12] presented three policies for load assignment among unequal strength components and compared three of these policies. Ye et al. [18] developed a model for a load sharing system where an operator dispatches work load to components in a manner that manages their degradation. He assumed degradation is the dominant failure type, and that the system will not be subject to sudden failure due to a shock. Yang et al. [19] proposed a novel approach for assessing a systems' reliability with dependency structures, load sharing, and damage accumulation and reversal. Zhao et al. [20] presented a reliability modeling and analysis framework for load-sharing systems with identical components subject to continuous degradation. Zhang et al. [21] proposed a new reliability analysis method for the load-sharing k-out-of-n: F system based on the load-strength model. Zhang et al. [22] presented a two-component load-sharing system. And the failure rates of the two components are time dependent and load dependent.

B. Motivation

Undoubtedly, above researches has contributed to the development of reliability of parallel system. Some of them propose excellent methods to calculate the mean time to failure of system, rest of them help to investigate the reliability of system under common cause failure or load-sharing failure respectively. However, lots of researches often ignore the joint of common cause and load-sharing failure in terms of the failure analysis of the parallel system with stress strength. Some parallel systems often happen simultaneously common cause and load-sharing failures. The results tend to be over positive than factual information. In this paper, for parallel systems, common cause failure model with stress strength and joint failure model of load-sharing and common cause failures are established respectively. Based on these models, the results are more approaching to the realistic situation considering the mean time to failure of parallel systems under common cause and load-sharing failures.

The rest of this paper is organized as follows. Section 2 analyzes the reliability model with stress strength under common cause failure. The reliability model under common cause and load-sharing failure is presented in section 3. Section 4 utilizes a numerical example to testify the validity of the proposed model. Finally, the conclusions of this paper are given in Section 5.

2. Reliability analysis with stress strength under common cause failure

Generally, x and y denote stress and strength respectively, $f_x(x)$ and $f_y(y)$ denote stress probability density function and strength probability density function respectively. We suppose parallel system is composed of n components. The probability of all components failure in the system is system conditional failure probability, so statistical average of system conditional failure probability under common cause failure is $p_s^n = \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^n dx$ where $x \in (0,+\infty)$. We utilize the model to calculate conditional failure probability of two-dependent-component and three-dependent-component parallel system respectively, and compare ultimate consequence.

According to above model, conditional failure probability of two-dependent-component parallel system is $p_s^2 = \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^2 dx$, and conditional failure probability of three-dependent-component parallel system is $p_s^3 = \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^3 dx$. Now we need to compare them. Because $\int_0^{+\infty} f_y(y)dy = 1$, we could get $0 \le \int_0^x f_y(y)dy \le 1$. $\int_0^x f_y(y)dy = 1$ when y is not more than x forever. However, y could be more than x. Thus, $0 \le \int_0^x f_y(y)dy < 1$. Based on relative mathematical knowledge, we can get $p_s^3 < p_s^2$. Obviously, conditional failure probability of three-dependent-component parallel system under common cause failure is less than two-dependent-component parallel system, which shows that we could decrease system conditional failure probability by increasing a redundant component.

For a parallel system under common cause failure with n components, if statistical average of system conditional failure probability satisfies $p_s^n = \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^n dx$, where x denotes stress, three properties could be deduced.

Property 1: conditional failure probability of three-dependent-component parallel system under common cause failure is less than two-dependent-component parallel system, where $x \in (0,+\infty)$.

Proof: $p_s^2 = \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^2 dx$, $p_s^3 = \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^3 dx$. Because $0 < f_x(x)\left[\int_0^x f_y(y)dy\right]^3 < f_x(x)\left[\int_0^x f_y(y)dy\right]^2$, based on relative mathematical knowledge, we could deduce $\int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^3 dx < \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^2 dx$. That is $p_s^3 < p_s^2$.

Property 2: conditional failure probability of k-dependent-component parallel system under common cause failure is less than (k-1)-dependent-component parallel system, where $x \in (0,+\infty)$.

Proof: $p_s^{k-1} = \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^{k-1} dx$,

$p_s^k = \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^k dx$. According to property 1, we can deduce $p_s^k < p_s^{k-1}$.

Property 3: system conditional failure probability approaches 0 when n approaches infinity, that is to say, $\lim_{n\to+\infty} p_s^n \approx 0$.

Proof: $p_s^n = \int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^n dx$, because $\int_0^x f_y(y)dy < 1$, $f_x(x)\left[\int_0^x f_y(y)dy\right]^n$ approaches infinitesimal when n approaches infinity. Thus, $\int_0^{+\infty} f_x(x)\left[\int_0^x f_y(y)dy\right]^n dx \approx 0$, that is to say, $\lim_{n\to+\infty} p_s^n \approx 0$.

Besides, $F_x(x)$ and $F_y(y)$ denote stress distribution function and strength distribution function. We suppose parallel system is composed of n components. Reliability of parallel system under common cause failure is $R_s = \int_{-\infty}^{+\infty}\left\{1-\left[\int_{-\infty}^x f_y(y)dy\right]^n\right\} f_x(x)dx = \int_{-\infty}^{+\infty}\left\{1-\left[F_y(x)\right]^n\right\} f_x(x)dx$ [13], where $x \in (0,+\infty)$.

According to above model, reliability of two-dependent-component parallel system is $R_s(2) = \int_{-\infty}^{+\infty}\left\{1-\left[F_y(x)\right]^2\right\} f_x(x)dx$, and reliability of three-dependent-component parallel system is $R_s(3) = \int_{-\infty}^{+\infty}\left\{1-\left[F_y(x)\right]^3\right\} f_x(x)dx$. Now we need to compare them.

$0 \le F_y(x) \le 1$, $F_y(x) = 1$ when y is not more than x forever. However, y could be more than x. Thus, $0 \le F_y(x) < 1$, according to above analysis, we can deduce $\left\{1-\left[F_y(x)\right]^3\right\} f_x(x) > \left\{1-\left[F_y(x)\right]^2\right\} f_x(x)$. Based on relative mathematical knowledge, we could deduce $R_s(2) < R_s(3)$. That is to say, reliability of three-dependent-component parallel system is more than two-dependent-component parallel system. Thus, we could deduce that increasing a redundant component would enhance system reliability.

For a parallel system under common cause failure with n components, if system reliability satisfies:

$$R_s = \int_{-\infty}^{+\infty} \left\{1-\left[\int_{-\infty}^{x} f_y(y)dy\right]^n\right\} f_x(x)dx = \int_{-\infty}^{+\infty} \left\{1-\left[F_y(x)\right]^n\right\} f_x(x)dx ,$$

where x and y denote stress and strength respectively, three properties could be deduced.

Property 4: Reliability of three-dependent-component parallel system under common cause failure is more than two-dependent-component parallel system, where $x \in (0,+\infty)$.

Proof: $R_s(2) = \int_{-\infty}^{+\infty} \left\{1-\left[F_y(x)\right]^2\right\} f_x(x)dx$,

$R_s(3) = \int_{-\infty}^{+\infty} \left\{1-\left[F_y(x)\right]^3\right\} f_x(x)dx$. Because $F_y(x) < 1$, we could get $\left[F_y(x)\right]^3 < \left[F_y(x)\right]^2$, and we could to deduce $1-\left[F_y(x)\right]^2 < 1-\left[F_y(x)\right]^3$. Thus, we could take a further step to deduce $\left\{1-\left[F_y(x)\right]^2\right\} f_x(x) < \left\{1-\left[F_y(x)\right]^3\right\} f_x(x)$. Based on relative mathematical knowledge, we get $R_s(2) < R_s(3)$.

Property 5: Reliability of k-dependent-component parallel system under common cause failure is more than (k-1)-dependent-component parallel system, where $x \in (0,+\infty)$.

Proof: $R_s(k-1) = \int_{-\infty}^{+\infty} \left\{1-\left[F_y(x)\right]^{k-1}\right\} f_x(x)dx$,

$R_s(k) = \int_{-\infty}^{+\infty} \left\{1-\left[F_y(x)\right]^{k}\right\} f_x(x)dx$. According to derivation way of property 4, we could deduce $R_s(k-1) < R_s(k)$.

Property 6: Parallel system reliability approaches 1 when n approaches infinity, that is to say, $\lim_{n\to+\infty} R_s(n) \approx 1$.

Proof: $R_s(n) = \int_{-\infty}^{+\infty} \left\{1-\left[F_y(x)\right]^{n}\right\} f_x(x)dx$, $\left[F_y(x)\right]^n < 1$. Thus, $1-\left[F_y(x)\right]^n \approx 1$ and $R_s(n) = \int_{-\infty}^{+\infty} \left\{1-\left[F_y(x)\right]^{n}\right\} f_x(x)dx \approx \int_{-\infty}^{+\infty} f_x(x)dx = 1$ when n approaches infinity.

3. Reliability analysis under load-sharing and common cause failures

We assume a system is composed of three same components. All components share whole system load and failure rate of each component is λ_3, when system works normally. Failure rate will become λ_2 with one component failed. When two components fail, failure rate will become λ_1. When there is one component working in the system, the common cause failure rate is λ_{c1}, when there are two components working in the system, the common cause failure rate is λ_{c2}, and when all of the three components are working normally, the common cause failure rate is λ_{c3}. We have merely one maintenance device which repairs randomly one failed component once, and other failed components must wait until last one has worked normally. With one component failed, μ_3 denote mean time to maintenance and maintenance rate respectively. With two components failed, μ_2 denote mean time to maintenance and maintenance rate respectively. With three components failed, μ_1 denote mean time to maintenance and maintenance rate respectively. According to the above assumption, we can describe the state transition figure of three-dependent-component parallel system under common cause and load-sharing failure as Fig. 1.

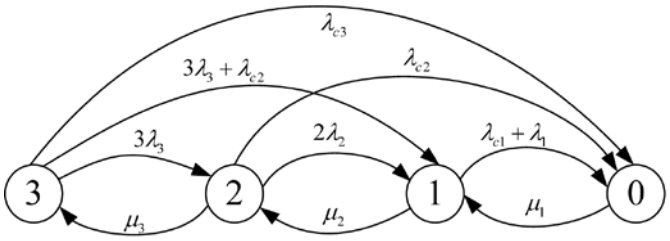


Fig. 1. state transition under common cause and load-sharing failures

As is shown in Fig. 1, based on state transition figure, we establish transition intensity matrix for calculation of system mean time to failure, and A denotes transition intensity matrix:

$$A = \begin{pmatrix} -\mu_1 & \mu_1 & 0 & 0 \\ \lambda_{c1}+\lambda_1 & -(\lambda_{c1}+\lambda_1+\mu_2) & \mu_2 & 0 \\ \lambda_{c2} & 2\lambda_2 & -(\lambda_{c2}+2\lambda_2+\mu_3) & \mu_3 \\ \lambda_{c3} & 3\lambda_3+\lambda_{c2} & 3\lambda_3 & -(6\lambda_3+\lambda_{c2}+\lambda_{c3}) \end{pmatrix}$$

The state 0 is absorbing state, therefore, we need to omit all elements in the system that is related to state 0. And B denotes a transition intensity matrix:

$$B = \begin{pmatrix} -(\lambda_{c1}+\lambda_1+\mu_2) & \mu_2 & 0 \\ 2\lambda_2 & -(\lambda_{c2}+2\lambda_2+\mu_3) & \mu_3 \\ 3\lambda_3+\lambda_{c2} & 3\lambda_3 & -(6\lambda_3+\lambda_{c2}+\lambda_{c3}) \end{pmatrix}$$

We have $C = \left[q_1(0) \quad q_2(0) \quad q_3(0)\right]$, $D = \left[0 \quad 0 \quad -1\right]$, where state transition equation is $CB = D$. Therefore, we could get the following equation:

$$\left[q_1(0) \quad q_2(0) \quad q_3(0)\right] \begin{pmatrix} -(\lambda_{c1}+\lambda_1+\mu_2) & \mu_2 & 0 \\ 2\lambda_2 & -(\lambda_{c2}+2\lambda_2+\mu_3) & \mu_3 \\ 3\lambda_3+\lambda_{c2} & 3\lambda_3 & -(6\lambda_3+\lambda_{c2}+\lambda_{c3}) \end{pmatrix} = \left[0 \quad 0 \quad -1\right]$$

Considering the complexity of equation and the accuracy of calculation, we can get $q_1(0), q_2(0), q_3(0)$ by using the math software. Then the mean time to failure of three-dependent-component parallel system is $\mathrm{MTTF_s}(3) = q_1(0) + q_2(0) + q_3(0)$. But the solution is too complex, we cannot use it to get some useful message, so we should make some assumptions to simplify the solution.

Assumptions 1: No matter how many components are working in the system, λ_c denotes common cause failure rate.

Assumptions 2: The failure rate decrease linearly with the decline of the quantity of the components which are working in the system, this is, if $\lambda_1=\lambda_e$, than we will get $\lambda_2=2\lambda_e$ and $\lambda_3=3\lambda_e$.

Assumptions 3: No matter how many components are broken, the maintenance rate is common, it is μ_1 .

Than we can get the simple version of the solution as follows:

$$\mathrm{MTTF_s}(3)-\mathrm{MTTF_s}(2)$$
$$=\frac{-\left(\lambda_c^2+7\lambda_c\lambda_e+5\lambda_c+12\lambda_e^2+9\lambda_e+6\lambda_e^2\right)\left(3\lambda_c\lambda_e-3\lambda_e^2\right)}{\left(\lambda_c^2+7\lambda_c\lambda_e+2\lambda_c+12\lambda_e^2\right)\left(2\lambda_c^3+20\lambda_c^2\lambda_e+10\lambda_c^2\lambda+66\lambda_c\lambda_e^2+39\lambda_c\lambda_e+6\lambda_c\lambda^2+72\lambda_e^3+27\lambda_e^2\lambda\right)}$$

However, the difference between the mean time to failure of four-dependent-component parallel system and three-dependent-component parallel system is more complex, even if it has been simplified, so it is hardly to find the same regular. Through the assumption we have made, we also can simplify the result of $q_1(0)$, $q_2(0)$, $q_3(0)$, they are:

$$q_1(0)=\frac{\lambda_c^2+7\lambda_c\lambda_e+3\mu\lambda_c+24\lambda e^2+9\mu\lambda_e}{2\lambda_c^3+20\lambda_c^2\lambda_e+8\lambda_c^2\mu+66\lambda_c\lambda_e^2+25\mu\lambda_c\lambda_e+72\lambda_e^3+3\lambda_e^2\mu-9\lambda_e\mu^2}$$

$$q_2(0)=\frac{3\lambda_c\lambda_e+2\mu\lambda_c+9\mu\lambda_e+9\lambda_e^2}{2\lambda_c^3+20\lambda_c^2\lambda_e+8\lambda_c^2\mu+66\lambda_c\lambda_e^2+25\mu\lambda_c\lambda_e+72\lambda_e^3+3\lambda_e^2\mu-9\lambda_e\mu^2}$$

$$q_3(0)=\frac{\lambda_c^2+7\lambda_c\lambda_e+4\mu\lambda_c+12\lambda_e^2+5\mu\lambda_e+3\mu^2}{2\lambda_c^3+20\lambda_c^2\lambda_e+8\lambda_c^2\mu+66\lambda_c\lambda_e^2+25\mu\lambda_c\lambda_e+72\lambda_e^3+3\lambda_e^2\mu-9\lambda_e\mu^2}$$

And then we can analyze the rate's influence of the component of the mean time to failure of three-dependent-component parallel system.

Firstly, we focus on the influence of λ_e to these components. As the denominator of the three components are same, when λ_e changes, there are the same changes happen in those denominators, so it is ok for us that do not care about the denominators, the only thing we should do is focusing on the numerator. We set the numerator of $q_1(0)$ is $Y_{1(\lambda_e)}$, taking the derivative of this function, we can get $Y_{1(\lambda_e)}'=7\lambda_c+9\mu+48\lambda_e$, similarly we can get the derivative of the numerator of $q_2(0)$ and $q_3(0)$, it is $Y_{2(\lambda_e)}'=3\lambda_c+9\mu+18\lambda_e$ and $Y_{3(\lambda_e)}'=7\lambda_c+5\mu+24\lambda_e$, and we know that these rates are positive, thus it is obviously that $Y_{1(\lambda_e)}'$ is the most through the three.

Secondly, analyzing the influence of λ_c , similarly we should focus on the numerator only. We set the numerator of $q_1(0)$ is $Y_{1(\lambda_c)}$, taking the derivative of this function, we can get $Y_{1(\lambda_c)}'=7\lambda_e+3\mu+2\lambda_c$, in the same way we can get the derivative of the numerator of $q_2(0)$ and $q_3(0)$, it is $Y_{2(\lambda_c)}'=3\lambda_e+2\mu$ and $Y_{3(\lambda_c)}'=7\lambda_e+4\mu+2\lambda_c$, we can find that $Y_{3(\lambda_c)}'>Y_{1(\lambda_c)}'>Y_{2(\lambda_c)}'$.

Finally, focusing on μ , also, the numerator is the only thing that we should care about. Setting the numerator of $q_1(0)$ is $Y_{1(\mu)}$, taking the derivative of this function, we can get $Y_{1(\mu)}'=3\lambda_c+9\lambda_e$, similarly we can get the derivative of the numerator of $q_2(0)$ and $q_3(0)$, it is $Y_{2(\mu)}'=2\lambda_c+9\lambda_e$ and $Y_{3(\mu)}'=4\lambda_c+5\lambda_e+6\mu$. But, as we cannot ensure the relative size of the rates, it is $Y_{1(\mu)}'>Y_{2(\mu)}'$ that we can find only.

Considering the above analysis, we can get some properties as follows.

Property 7: The failure rate λ_e has the most influence on the system when there is only a component working in.

Proof: Under the premise that all of the three rates are positive, consider the derivatives above, $Y_{1(\lambda_e)}'=7\lambda_c+9\mu+48\lambda_e$, $Y_{2(\lambda_e)}'=3\lambda_c+9\mu+18\lambda_e$ and $Y_{3(\lambda_e)}'=7\lambda_c+5\mu+24\lambda_e$, we can find that $Y_{1(\lambda_e)}'$ is the most one, so we can say the failure rate λ_e has the most influence to the system when there is only a component working.

Property 8: The common cause failure rate λ_c has the most influence on the system when there are three components working in, has the second most influence on it when there are two components working in, and has the least influence on the system when there is only a component working in.

Proof: Under the premise that all of the three rates are positive, consider the derivatives above, $Y_{1(\lambda_c)}'=7\lambda_e+3\mu+2\lambda_c$, $Y_{2(\lambda_c)}'=3\lambda_e+2\mu$ and $Y_{3(\lambda_c)}'=7\lambda_e+4\mu+2\lambda_c$, we can find that $Y_{3(\lambda_c)}'>Y_{1(\lambda_c)}'>Y_{2(\lambda_c)}'$. Therefore, we can get the view that the common cause failure rate λ_c has the most influence on the system when there are three components working in, has the second most influence on it when there are two components working in, and has the least influence on the system when there is only a component working in.

Property 9: The maintenance rate μ influences the system when there is only a component working in more than when there are two components working in.

Proof: Under the premise that all of the three rates are positive, consider the derivatives above, $Y_{1(\mu)}'=3\lambda_c+9\lambda_e$, $Y_{2(\mu)}'=2\lambda_c+9\lambda_e$ and $Y_{3(\mu)}'=4\lambda_c+5\lambda_e+6\mu$, we can easily find that $Y_{1(\mu)}'>Y_{2(\mu)}'$, so we can say the maintenance rate μ influences the system when there is only a component working in more than when there are two components working in.

4. Numerical example

In this section, we will have an analysis about a parallel system of three components under common cause and load-sharing failure. This section mainly studies the effect of single variance on the reliability of parallel system. We assume the reliability parameters are $\lambda_a=1.12\times10^{-3}h^{-1}$, $\lambda_m=2\times10^{-3}h^{-1}$, $\lambda_f=5\times10^{-3}h^{-1}$, $\lambda_c=3\times10^{-4}h^{-1}$, $MDT_a=16h$, $\mu_a=5.2\times10^{-3}h^{-1}$, $MDT_k=12h$, $\mu_k=8.3\times10^{-2}h^{-1}$, $MDT_f=8h$, $\mu_f=1.25\times10^{-1}h^{-1}$.

4.1. The effect of each failure rate on mean time to failure of parallel system

(1) Effect of λ_a variation on $\mathrm{MTTF_s}$

λ_a is defined as independent variable, and its range of values is $\left[0,2\times10^{-3}\right]$. Dependent variable is $\mathrm{MTTF_s}$. We can calculate mean time to failure of two-dependent-component and three-dependent-component parallel system under common cause and load-sharing failure.

$$\mathrm{MTTF_s}(2)=p_1(0)+p_2(0)=\frac{\lambda_c+\lambda_f+2\lambda_m+\mu_k}{(\lambda_c+2\lambda_m)(\lambda_c+\lambda_f+\mu_k)-2\lambda_m\mu_k}=1935.4(h)$$

$$\mathrm{MTTF_s}(3)=q_1(0)+q_2(0)+q_3(0)=$$
$$\frac{6\lambda_a\lambda_c+3\lambda_a\lambda_f+12\lambda_a\lambda_m+2\lambda_c^2+\lambda_c\lambda_f+4\lambda_c\lambda_m+2\lambda_f\lambda_m+3\lambda_a\mu_a+6\lambda_a\mu_k+2\lambda_c\mu_a+2\lambda_c\mu_k+\lambda_f\mu_a+\mu_a\mu_k}{(\lambda_c+\lambda_f+\mu_k)(\lambda_c+2\lambda_m+\mu_a)(6\lambda_a+2\lambda_c)-3\lambda_a\mu_a(\lambda_c+\lambda_f+2\mu_k)-\mu_k(12\lambda_a\lambda_m+4\lambda_c\lambda_m+\lambda_c\mu_a)}=$$
$$\frac{0.5544\lambda_a+5.346\times10^{-4}}{3.6882\times10^{-4}\lambda_a+1.7463\times10^{-7}}$$

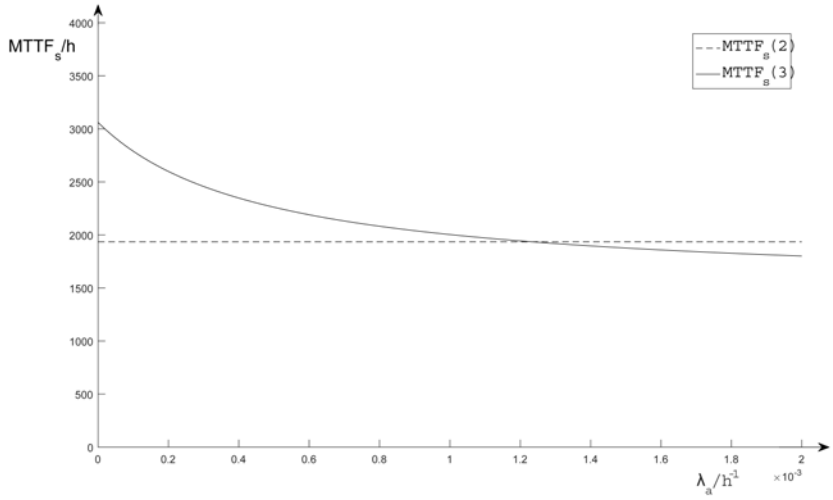


Fig. 2. Effect of λ_a *variation on* $\mathrm{MTTF_s}$

Fig. 2 describes the effect λ_a variation on $\mathrm{MTTF_s}(2)$ and $\mathrm{MTTF_s}(3)$. Firstly, $\mathrm{MTTF_s}(2)$=1935.4(h), λ_a have no effect on mean time to failure of two-dependent-component parallel system. Secondly, mean time to failure of three-dependent-component parallel system is negatively correlated with λ_a. Thirdly, $\begin{cases} \mathrm{MTTF_s}(3)>\mathrm{MTTF_s}(2), 0 \le \lambda_a < 1.233\times10^{-3} \\ \mathrm{MTTF_s}(3)=\mathrm{MTTF_s}(2), \lambda_a = 1.233\times10^{-3} \\ \mathrm{MTTF_s}(3)<\mathrm{MTTF_s}(2), 1.233\times10^{-3} < \lambda_a \le 2\times10^{-3} \end{cases}$, so the three-dependent-component parallel system is prior to two-dependent-component parallel system when $0 \le \lambda_a < 1.233\times10^{-3}$.

(2) The effect of λ_m variation on $\mathrm{MTTF_s}$

λ_m is defined as independent variable, and its range of values is $\left[0, 5\times10^{-3}\right]$. Dependent variable is $\mathrm{MTTF_s}$. We can calculate mean time to failure of two-dependent-component and three-dependent-component parallel system under common cause and load-sharing failure.

$$\mathrm{MTTF_s}(2) = p_1(0) + p_2(0) = \frac{\lambda_c + \lambda_f + 2\lambda_m + \mu_k}{(\lambda_c + 2\lambda_m)(\lambda_c + \lambda_f + \mu_k) - 2\lambda_m\mu_k} = \frac{2\lambda_m + 8.830\times10^{-2}}{1.060\times10^{-2}\lambda_m + 2.649\times10^{-5}}$$

$$\mathrm{MTTF_s}(3) = q_1(0) + q_2(0) + q_3(0) =$$
$$\frac{6\lambda_a\lambda_c + 3\lambda_a\lambda_f + 12\lambda_a\lambda_m + 2\lambda_c^2 + \lambda_c\lambda_f + 4\lambda_c\lambda_m + 2\lambda_f\lambda_m + 3\lambda_a\mu_a + 6\lambda_a\mu_k + 2\lambda_c\mu_a + 2\lambda_c\mu_k + \lambda_f\mu_a + \mu_a\mu_k}{(\lambda_c + \lambda_f + \mu_k)(\lambda_c + 2\lambda_m + \mu_a)(6\lambda_a + 2\lambda_c) - 3\lambda_a\mu_a(\lambda_c + \lambda_f + 2\mu_k) - \mu_k(12\lambda_a\lambda_m + 4\lambda_c\lambda_m + \lambda_c\mu_a)} =$$
$$\frac{2.464\lambda_m + 1.106\times10^{-1}}{7.759\times10^{-3}\lambda_m + 4.325\times10^{-5}}$$

Now we describe the effect λ_m variation on $\mathrm{MTTF_s}(2)$ and $\mathrm{MTTF_s}(3)$ in Fig. 3.

Firstly, mean time to failure of three-dependent-component and two-dependent-component parallel system is negatively correlated with λ_m. Secondly, $\begin{cases} \mathrm{MTTF_s}(3)<\mathrm{MTTF_s}(2), 0 \le \lambda_m < 1.832\times10^{-3} \\ \mathrm{MTTF_s}(3)=\mathrm{MTTF_s}(2), \lambda_m = 1.832\times10^{-3} \\ \mathrm{MTTF_s}(3)>\mathrm{MTTF_s}(2), 1.832\times10^{-3} < \lambda_m \le 5\times10^{-3} \end{cases}$, so the three-dependent-component parallel system is prior to two-dependent-component parallel system when $\lambda_m \in \left[1.832\times10^{-3}, 5\times10^{-3}\right]$.

(3)The effect of λ_f variation on $\mathrm{MTTF_s}$

λ_f is defined as independent variable, and its range of values is $\left[2\times10^{-3}, 1\times10^{-2}\right]$.

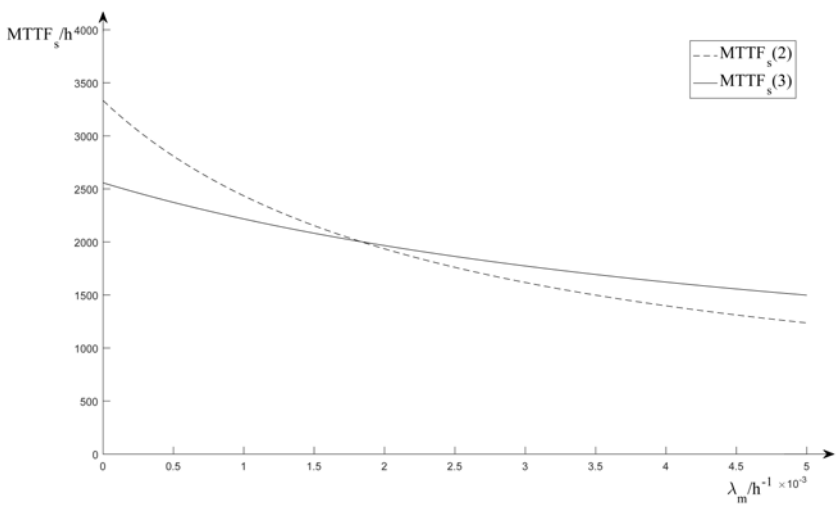


Fig. 3. Effect of λ_m *variation on* $\mathrm{MTTF_s}$

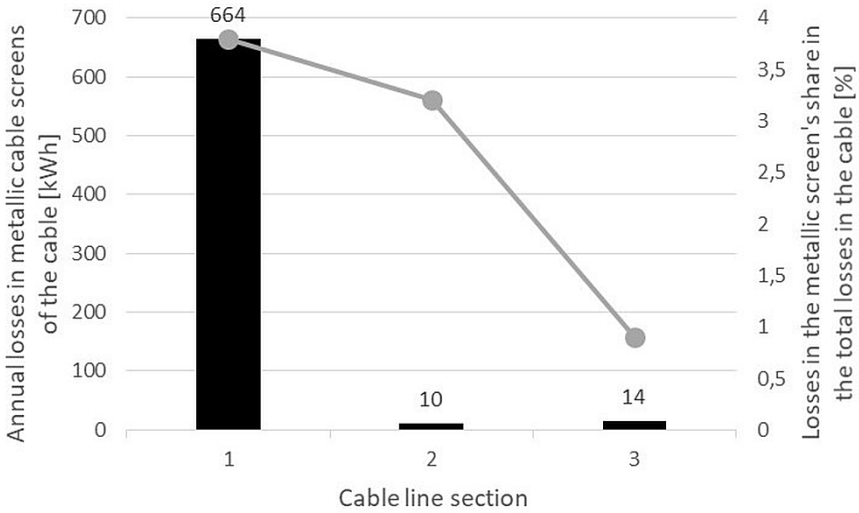


Fig. 4. Effect of λ_f *variation on* $\mathrm{MTTF_s}$

$$\mathrm{MTTF_s}(2)=p_1(0)+p_2(0)=\frac{\lambda_c+\lambda_f+2\lambda_m+\mu_k}{(\lambda_c+2\lambda_m)(\lambda_c+\lambda_f+\mu_k)-2\lambda_m\mu_k}=\frac{\lambda_f+8.730\times10^{-2}}{4.300\times10^{-3}\lambda_f+2.619\times10^{-5}}$$

$$\mathrm{MTTF_s}(3)=q_1(0)+q_2(0)+q_3(0)=$$
$$\frac{6\lambda_a\lambda_c+3\lambda_a\lambda_f+12\lambda_a\lambda_m+2\lambda_c^2+\lambda_c\lambda_f+4\lambda_c\lambda_m+2\lambda_f\lambda_m+3\lambda_a\mu_a+6\lambda_a\mu_k+2\lambda_c\mu_a+2\lambda_c\mu_k+\lambda_f\mu_a+\mu_a\mu_k}{(\lambda_c+\lambda_f+\mu_k)(\lambda_c+2\lambda_m+\mu_a)(6\lambda_a+2\lambda_c)-3\lambda_a\mu_a(\lambda_c+\lambda_f+2\mu_k)-\mu_k(12\lambda_a\lambda_m+4\lambda_c\lambda_m+\lambda_c\mu_a)}=$$
$$\frac{1.286\lambda_f+1.091\times10^{-1}}{5.207\times10^{-3}\lambda_f+3.273\times10^{-5}}$$

As is shown in Fig. 4, the three-dependent-component parallel system is prior to two-dependent-component parallel system when $\lambda_f\in\left[2\times10^{-3},1\times10^{-2}\right]$.

(4)The effect of λ_c variation on $\mathrm{MTTF_s}$

λ_c is defined as independent variable, and its range of values is $\left[0,1\times10^{-3}\right]$. Dependent variable is $\mathrm{MTTF_s}$. We can calculate mean time to failure of two-dependent-component and three-dependent-component parallel system under common cause and load-sharing failure:

$$\mathrm{MTTF_s}(2)=p_1(0)+p_2(0)=\frac{\lambda_c+\lambda_f+2\lambda_m+\mu_k}{(\lambda_c+2\lambda_m)(\lambda_c+\lambda_f+\mu_k)-2\lambda_m\mu_k}=\frac{(\lambda_c+9.200\times10^{-2})}{(\lambda_c+8.800\times10^{-2})(\lambda_c+4\times10^{-3})-3.32\times10^{-4}}$$

$$\mathrm{MTTF_s}(3)=q_1(0)+q_2(0)+q_3(0)=$$
$$\frac{6\lambda_a\lambda_c+3\lambda_a\lambda_f+12\lambda_a\lambda_m+2\lambda_c^2+\lambda_c\lambda_f+4\lambda_c\lambda_m+2\lambda_f\lambda_m+3\lambda_a\mu_a+6\lambda_a\mu_k+2\lambda_c\mu_a+2\lambda_c\mu_k+\lambda_f\mu_a+\mu_a\mu_k}{(\lambda_c+\lambda_f+\mu_k)(\lambda_c+2\lambda_m+\mu_a)(6\lambda_a+2\lambda_c)-3\lambda_a\mu_a(\lambda_c+\lambda_f+2\mu_k)-\mu_k(12\lambda_a\lambda_m+4\lambda_c\lambda_m+\lambda_c\mu_a)}=$$
$$\frac{2\lambda_c^2+0.196\lambda_c+1.097\times10^{-3}}{(2\lambda_c+6.72\times10^{-3})(\lambda_c+8.8\times10^{-2})(\lambda_c+9.2\times10^{-3})-1.113\times10^{-3}\lambda_c-5.219\times10^{-6}}$$

Now we describe the figure of the effect λ_c variation on $\mathrm{MTTF_s}(2)$ and $\mathrm{MTTF_s}(3)$ in Fig. 5.

Firstly, mean time to failure of three-dependent-component and two-dependent-component parallel system is negatively correlated with λ_m. Secondly, $\begin{cases}\mathrm{MTTF_s}(3)>\mathrm{MTTF_s}(2), 0\le\lambda_c<5\times10^{-4}\\ \mathrm{MTTF_s}(3)=\mathrm{MTTF_s}(2), 5\times10^{-4}\le\lambda_c\le1\times10^{-3}\end{cases}$, so the three-dependent-component parallel system is prior to two-dependent-component parallel system when $\lambda_c\in\left[0,5\times10^{-4}\right]$.

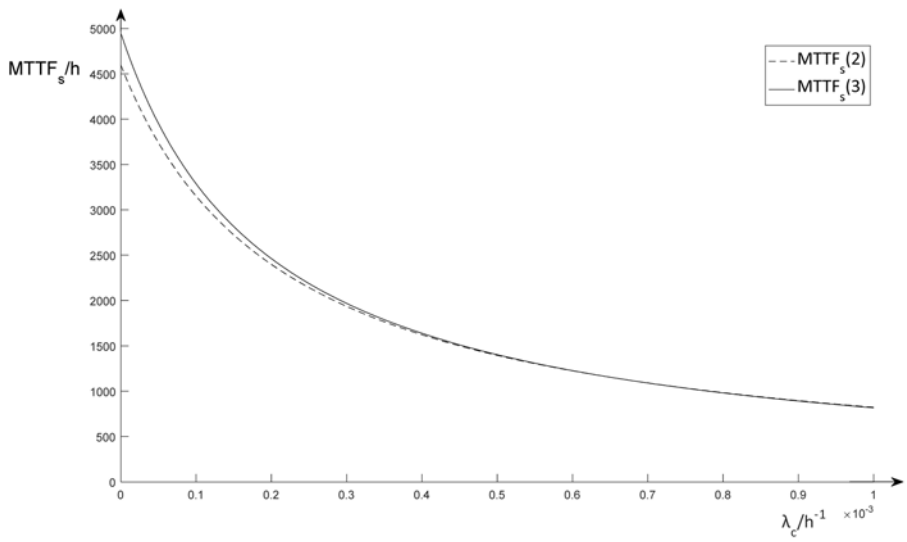


Fig. 5. Effect of λ_c *variation on* $\mathrm{MTTF_s}$

4.2. The effect of each maintenance rate on mean time to failure of parallel system

(1)The effect of μ_a variation on $\mathrm{MTTF_s}$

μ_a is defined as independent variable, and its range of values is $\left[0, 8.3\times10^{-2}\right]$. Dependent variable is $\mathrm{MTTF_s}$. We can calculate mean time to failure of two-dependent-component and three-dependent-component parallel system under common cause and load-sharing failure.

$$\mathrm{MTTF_s}(2) = p_1(0) + p_2(0) = \frac{\lambda_c + \lambda_f + 2\lambda_m + \mu_k}{(\lambda_c + 2\lambda_m)(\lambda_c + \lambda_f + \mu_k) - 2\lambda_m\mu_k} \approx 1935.4(h),$$

$$\mathrm{MTTF_s}(3) = q_1(0) + q_2(0) + q_3(0) = \frac{9.196\mu_a + 6.773\times10^{-2}}{4.589\times10^{-3}\mu_a + 3.491\times10^{-5}}.$$

In Fig. 6, the mean time to failure of three-dependent-component parallel system is weakly positive correlation with μ_a. When $\mu_a=0$, $\mathrm{MTTF_s}(3)=\mathrm{MTTF_s}(2)$. The three-dependent-component parallel system is prior to two-dependent-component parallel system when $\mu_a \in \left[0, 8.3\times10^{-2}\right]$.

(2)The effect of μ_k variation on $\mathrm{MTTF_s}$

μ_k is defined as independent variable, and its range of values is $\left[5.2\times10^{-3}, 0.125\right]$. Dependent variable is $\mathrm{MTTF_s}$.

$$\mathrm{MTTF_s}(2) = p_1(0) + p_2(0) = \frac{\lambda_c + \lambda_f + 2\lambda_m + \mu_k}{(\lambda_c + 2\lambda_m)(\lambda_c + \lambda_f + \mu_k) - 2\lambda_m\mu_k} = \frac{\mu_k + 9.3\times10^{-3}}{3\times10^{-4}\mu_k + 2.279\times10^{-5}},$$

$$\mathrm{MTTF_s}(3) = q_1(0) + q_2(0) + q_3(0) =$$

$$\frac{6\lambda_a\lambda_c + 3\lambda_a\lambda_f + 12\lambda_a\lambda_m + 2\lambda_c^2 + \lambda_c\lambda_f + 4\lambda_c\lambda_m + 2\lambda_f\lambda_m + 3\lambda_a\mu_a + 6\lambda_a\mu_k + 2\lambda_c\mu_a + 2\lambda_c\mu_k + \lambda_f\mu_a + \mu_a\mu_k}{(\lambda_c + \lambda_f + \mu_k)(\lambda_c + 2\lambda_m + \mu_a)(6\lambda_a + 2\lambda_c) - 3\lambda_a\mu_a(\lambda_c + \lambda_f + 2\mu_k) - \mu_k(12\lambda_a\lambda_m + 4\lambda_c\lambda_m + \lambda_c\mu_a)} =$$

$$\frac{1.252\mu_k + 1.164\times10^{-2}}{3.756\times10^{-4}\mu_k + 2.760\times10^{-5}}$$

From Fig. 7, the mean time to failure of three-dependent-component and two-dependent-component parallel system is positive correlation with μ_k. The three-dependent-component parallel system is

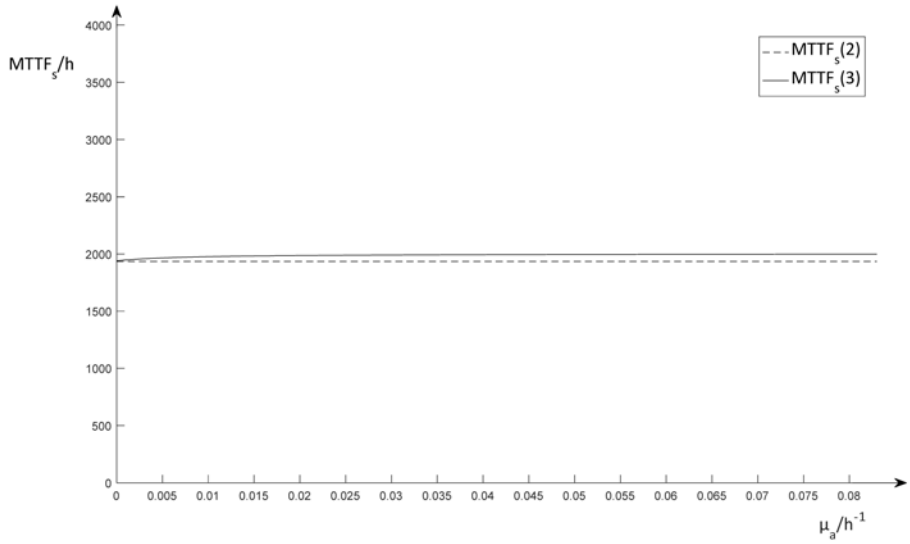


Fig. 6. Effect of μ_a variation on $\mathrm{MTTF_s}$

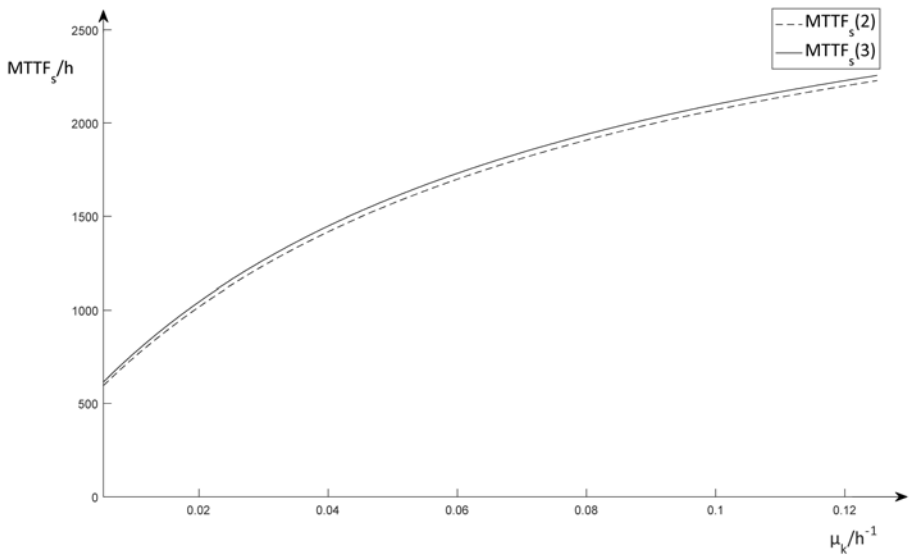


Fig. 7. Effect of variation on

prior to two-dependent-component parallel system when $\mu_k \in \left[5.2\times10^{-3}, 0.125\right]$.

5. Conclusions

This paper presents parallel system model under common cause and load-sharing failures. According to this model, mean time to failure of three-dependent-component and two-dependent-component parallel systems is calculated. Besides, we calculate and discuss the conditional failure probability and reliability of three-dependent-component and two-dependent-component parallel system model under common cause failure. The reliability of three-dependent-component parallel system model under common cause failure is more than two-dependent-component. We could observe that mean time to failure of three-dependent-component parallel systems is not always longer than two-dependent-component. Hence, some measures could be taken to control the range of variables to ensure mean time to failure of three-dependent-component parallel systems is more than two-dependent-component.

Acknowledgements

The authors gratefully acknowledge the financial supports for this research from the National Natural Science Foundation of China (Nos. 51875015, 51620105010, 51705526).

References

1. Alizadeh S, Srirramula S. Impact of common cause failure on reliability performance of redundant safety related systems subject to process demand. Reliability Engineering & System Safety 2018; 172: 129-150, https://doi.org/10.1016/j.ress.2017.12.011.
2. Fan M F, Zeng Z G. A stochastic hybrid systems model of common-cause failures of degrading components., Reliability Engineering & System Safety 2018; 172: 159-170, https://doi.org/10.1016/j.ress.2017.12.003.
3. Huang L, Xu Q. Lifetime reliability for load-sharing redundant systems with arbitrary failure distributions. IEEE Transactions on Reliability 2010; 59: 319-330, https://doi.org/10.1109/TR.2010.2048679.
4. Jiang T, Liu Y, Zheng Y. Optimal loading strategy for multi-state systems: cumulative performance perspective. Applied Mathematical Modelling 2019; 74: 199-216, https://doi.org/10.1016/j.apm.2019.04.043.
5. Levitin G. Incorporating common- cause failures into nonrepairable multistate series-parallel system analysis. IEEE Transactions on Reliability 2001; 50: 380-388, https://doi.org/10.1109/24.983398.
6. Liu H M. Reliability of a load-sharing k-out-of-n:G system: non-iid components with arbitrary distributions. IEEE Transactions on Reliability 1998; 47: 279-284, https://doi.org/10.1109/24.740502.
7. Pourali M. Incorporating common cause failures in mission-critical facilities reliability analysis. IEEE Transactions on Industry Applications 2014; 50: 2883-2890, https://doi.org/10.1109/TIA.2013.2295472.
8. Paula C P, Visnadi L B, Castro H F. Multi-objective optimization in redundant system considering load sharing. Reliability Engineering & System Safety 2019; 181: 17-27, https://doi.org/10.1016/j.ress.2018.08.012.
9. Sakurahare T, Schumock G, Reihani S. Simulation-informed probabilistic methodology for common cause failure analysis. Reliability Engineering & System Safety 2019; 185: 84-99, https://doi.org/10.1016/j.ress.2018.12.007.
10. Sutar S S, Naik-Nimbalkar U V. Accelerated failure time models for load sharing systems. IEEE Transactions on Reliability 2014; 63: 706-714, https://doi.org/10.1109/TR.2014.2313793.
11. Vaurio J K. An implicit method for incorporating common-cause failures in system analysis. IEEE Transactions on Reliability 1998; 47: 173-180, https://doi.org/10.1109/24.722285.
12. Wang K S, Huang J J, Tsai Y T, Hsu F S. Study of loading policies for unequal strength shared-load system. Reliability Engineering & System Safety 2000; 67: 119-128, https://doi.org/10.1016/S0951-8320(99)00057-5.
13. Wang C N, Xing L D, Levitin G. Probability common cause failures in phased-mission systems. Reliability Engineering & System Safety 2015; 144: 53-60, https://doi.org/10.1016/j.ress.2015.07.004
14. Wang C N, Xing L D, Levitin G. Explicit and implicit methods for probabilistic common-cause failure analysis. Reliability Engineering & System Safety 2014; 131: 175-184, https://doi.org/10.1016/j.ress.2014.06.024.
15. Xiao H, Gao S. Simulation budget allocation for selecting the top-m designs with input uncertainty. IEEE Transactions on Automatic Control 2018; 63: 3127-3134, https://doi.org/10.1109/TAC.2018.2791425.
16. Xiao H, Chen H, Lee L H. An efficient simulation procedure for ranking the top simulated designs in the presence of stochastic constraints. Automatica 2019; 103: 106-115, https://doi.org/10.1016/j.automatica.2018.12.008.
17. Yuan J. Pivotal decomposition to find availability and failure-frequency of systems with common-cause failures. IEEE Transactions on Reliability 1987; 36: 48-53, https://doi.org/10.1109/TR.1987.5222292
18. Ye Z S, Revie M, Walls L. A load sharing system reliability model with managed component degradation. IEEE Transactions on Reliability 2014; 63: 721-730, https://doi.org/10.1109/TR.2014.2315965.
19. Yang K, Younis H. A semi-analytical Monte Carlo simulation method for system's reliability with load sharing and damage accumulation. Reliability Engineering & System Safety 2005; 87: 191-200, https://doi.org/10.1016/j.ress.2004.04.016.
20. Zhao X J, Liu B, Liu Y Q. Reliability modeling and analysis of load-sharing systems with continuously degrading components. IEEE Transactions on Reliability 2018; 67: 1096-1110, https://doi.org/10.1109/TR.2018.2846649.
21. Zhang J C, Zhao Y, Ma X B. A new reliability analysis method for load-sharing k-out-of-n: F system based on load-strength model. Reliability Engineering & System Safety 2019; 182: 152-165, https://doi.org/10.1016/j.ress.2018.10.013.
22. Zhang N, Fouladirad M, Barros A. Maintenance analysis of a two-component load- sharing system. Reliability Engineering & System Safety 2017; 167: 67-74, https://doi.org/10.1016/j.ress.2017.05.027.
23. Zuo L, Tangfan X, Yu L. Evidential network-based failure analysis for systems suffering common cause failure and model parameter uncertainty. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 2019; 233: 2225-2235, https://doi.org/10.1177/0954406218781407.

Chao ZHANGA
School of Automation Science and Electrical Engineering
Beihang University
Beijing 100083, China

Yunan ZHANGB
Intelligence Science College
National University of Defense Technology
Changsha 410073, China

Email: zhangchaobuaa@126.com, dhynhnh@126.com

Article citation info:
MŁYNARSKI S, PILCH R, SMOLNIK M, SZYBKA J, WIĄZANIA G. A model of an adaptive strategy of preventive maintenance of complex technical objects. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 35–41, http://dx.doi.org/10.17531/ein.2020.1.5.

Stanisław MŁYNARSKI
Robert PILCH
Maksymilian SMOLNIK
Jan SZYBKA
Grzegorz WIĄZANIA

A MODEL OF AN ADAPTIVE STRATEGY OF PREVENTIVE MAINTENANCE OF COMPLEX TECHNICAL OBJECTS

MODEL ADAPTACYJNEJ STRATEGII PREWENCYJNEJ ODNOWY ZŁOŻONYCH OBIEKTÓW TECHNICZNYCH*

The paper presents results of the analysis of the developed models for complex technical objects preventive maintenance scheduling. Models based on two different sets of assumptions were developed. The general problem solved was to determine the joint time of preventive renewal for a group of parts or subassemblies. The purpose of the first model (the model of scheduled preventive maintenance strategy) is to determine the profitability of constant application of a previously developed preventive maintenance schedule for a part undergoing post-failure renewal. The second model (the model of adaptive strategy of a system's preventive maintenance) allows one to determine a new joint time of preventive renewal for a group of parts each time when one of them is undergoing a post-failure renewal. The initial preventive maintenance strategy for each part or subassembly was obtained using typical tools for maintenance planning (decision-random models based on dynamic programming and Bellman's principle of optimality). Exemplary simulation calculations with the use of both models were made and their results presented as the total maintenance costs estimated for the renewal strategies developed. The object of the analysis were the chosen geometrical features of a rail vehicle wheel changing due to its wear during operation. Based on this kind of analysis, one can choose a better preventive maintenance model for a specific application area.

***Keywords**: preventive maintenance; simulation efficiency evaluation; reliability engineering.*

W artykule przedstawiono wyniki analizy opracowanych modeli do planowania odnowy profilaktycznej złożonych obiektów technicznych, które oparto o dwa różniące się od siebie zestawy założeń. Rozwiązywany problem dotyczy określenia wspólnego czasu odnowy profilaktycznej grupy części lub podzespołów złożonego obiektu. Pierwszy z opracowanych modeli (model planowej strategii odnowy prewencyjnej) pozwala określić zasadność przeprowadzenia ustalonego wcześniej, planowego odnowienia prewencyjnego części obiektu, która została już odnowiona poawaryjnie. Drugi model (model adaptacyjnej strategii odnowy prewencyjnej) umożliwia wyznaczenie najbliższego wspólnego czasu odnowy profilaktycznej grupy części, z których jedna aktualnie podlega odnowie poawaryjnej. Początkowe (wyjściowe) strategie odnowy profilaktycznej każdej części bądź podzespołu wyznaczone zostały za pomocą standardowych narzędzi do planowania odnawiania profilaktycznego (modeli decyzyjno-losowych wykorzystujących programowanie dynamiczne Bellmana). Posługując się opracowanymi modelami odnowy, przeprowadzono przykładowe obliczenia symulacyjne, których wyniki przedstawiono w postaci całkowitych kosztów obsługiwania dla każdej z uzyskanych strategii. Przedmiotem analizy były wybrane cechy geometryczne koła pojazdu szynowego, których wartości zmieniają się na skutek zużycia w procesie eksploatacji. Na podstawie tego rodzaju analiz można wybrać lepszy (tj. efektywniejszy ekonomicznie) z modeli dla konkretnego zastosowania w praktyce.

***Słowa kluczowe**: odnowa profilaktyczna; symulacyjna ocena efektywności; inżynieria niezawodności.*

Notations

t_i^* – preventive maintenance time of i-th element of the system,
k_{ai} – failure cost of i-th element of the system,
k_{oi} – preventive maintenance cost of i-th element of the system,
t_{max}^* – the highest value of preventive maintenance time of one out of all the elements of the system,
t_{min}^* – the lowest value of preventive maintenance time of one out of all the elements of the system,
t_{ai} – time of failure of i-th element,
t^* – joint time to preventive maintenance of elements determined prior to failure of i th element,
$F_i(t)$ – cumulative distribution function of time to failure of i-th element,
k_{ai} – cost of failure of i-th element of the system,
k_{oi} – cost of preventive maintenance of i-th element of the system,
$D(t_{ai})$ – decision indicator describing the profitability of renewal,
T_h – finite time horizon of the object's operation,
T_{hmax} – maximal operation time without preventive renewal and/or inspection,
Δt – time interval in calculations with the use of decision-random models,

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

n_{OT} – number of maintenance activities in the adopted time horizon,
n – number of simulations,
dt – time step used during simulation,
t_U – age of the system (counted from the beginning of simulation),
t_{Ei} – age of the i-th element (counted from the beginning of simulation),
Pu_{Ei} – probability of failure of the i-th element in simulation calculations,
RND_{Ei} – random number from the interval [0,1], of uniform distribution, for the i-th element,
t_{O_opt} – system's optimal renewal time period in simulation calculations.

1. Introduction

To guarantee economically effective and safe use of complex technical objects it is necessary to employ an adequate strategy of their maintenance [4]. The quality of the operation process, described by various indicators [19], depends largely on the correct renewal of technical objects. This problem is of particular significance in the case of systems in which the failure of elements may threaten human health and life or result in considerable economic losses. In such context, preventive maintenance of system's elements may be justified, and the optimal time of its execution is determined by various models of preventive maintenance [15, 20]. To this end economic data on failures and maintenance as well as reliability characteristics of system's elements are used. In the literature, apart from the classical maintenance models [10], there are numerous maintenance models available on the required reliability level of an entire system [12, 21], some of them based on the application of simulation methods [5, 8]. Some models include the possibility of partial maintenance (cf. [7, 8]), some others make use of additional inspection of object's technical state when it can be performed while the system is in actual operation [2, 3]. A detailed and comprehensive classification of existing preventive renewal models is provided in [20].

A model for the determination of the preventive maintenance optimal time (t*) should be adapted to the specificity and operation conditions of the given object. Only then will the preventive maintenance strategy based on this model be economically justified and guarantee the reduced failure rate of the object (cf. [11, 14]).

The aim of the paper is to develop an adaptive model of renewal of complex technical objects and a simulation method to assess the efficiency of using the renewal strategy determined with this model.

The article discusses the specific problem of renewal of a complex technical object, whose parts are functioning in a reliability series structure and their failures are considered as independent of each other, however, for organizational and economic reasons, individual renewal of each of them may be less profitable than block renewal.

In complicated cases, when the maintenance of some parts of complex technical objects is performed at the same time as the servicing of other units or systems of the given object, after a failure the scheduled maintenance strategy of the failed part or even the entire object needs to be modified. The present paper is devoted to the analysis of this important practical problem.

The developed solutions are a significant complement to existing renewal models due to the needs in maintenance practice.

The article presents a new original computer simulation based method of the evaluation of the efficiency of the proposed strategies. The newly developed method was validated on a practical example of failures of wheels of an electric rail unit vehicle.

2. Models of preventive maintenance of complex systems

Technical objects are composed of many structural elements that have individual operational characteristics. The empirical reliability indexes of component parts are often described with various models and their plots differ considerably [7]. As a result, the decision making as to the dates and scope of the safety-orientated maintenance of entire objects or their units and parts, and, consequently, their rational and economically effective use is difficult (cf. [6, 16]).

2.1. Determination of joint preventive maintenance time of system's components

When preventive maintenance of systems that are composed of a number of different elements is scheduled, for organisational and economic reasons a joint time of the preventive maintenance of all the elements, that is the entire system, is most frequently searched for [17]. If the preventive maintenance model adopted includes a unit of the complex technical object considered as a whole – there is a reliability function as well as maintenance and failure cost specified for it, the indicated optimal maintenance time refers to this entire unit. Moreover, it is subject to preventive and post-failure maintenance always as a whole (all its parts are subject to maintenance simultaneously).

The problem of the determination of a joint preventive maintenance time arises when for particular elements of the system optimal (following the adopted preventive maintenance model) different from each other, but close, preventive maintenance times are determined. The proposed manner of unification of preventive maintenance time assumes that individual preventive maintenance times of component parts are accounted for (t_i^*), as well as preventive maintenance costs (k_{oi}), failure costs (k_{ai}) and increments of distribution functions as to the range of individual preventive maintenance times for all the elements. The distribution function increments, maintenance and failure costs will for each of the elements be the weight of its individual preventive maintenance time (t_i^*). We propose the joint preventive maintenance time (t^*) to be determined from the formula:

$$t^* = \frac{\sum_{i=1}^{n} t_i^* \cdot (k_{ai} + k_{oi}) \cdot \left[F_i\left(t_{max}^*\right) - F_i\left(t_{min}^*\right) \right]}{\sum_{i=1}^{n} (k_{ai} + k_{oi}) \cdot \left[F_i\left(t_{max}^*\right) - F_i\left(t_{min}^*\right) \right]} \tag{1}$$

where:
t_i^* – preventive maintenance time of i-th element of the system,
k_{ai} – failure cost of i-th element of the system,
k_{oi} – preventive maintenance cost of i-th element of the system,
t_{max}^* – the highest value of preventive maintenance time of one out of all the elements of the system,
t_{min}^* – the lowest value of preventive maintenance time of one out of all the elements of the system.

The joint preventive maintenance time of the given unit (t^*) can also be defined including scheduled maintenance cycles and servicing of other units and parts of the same complex object [9]. Such an approach is often followed in the operational practice of large transport companies. In these companies the optimal maintenance periods based on the adopted preventive maintenance model and inspections of individual units are extended or reduced so that they constitute an integer multiple of the scheduled maintenance activities of a larger number of units and parts of the object. Such an approach results in the reduction of downtime periods of a complex technical object, which brings measurable economic effects.

2.2. A model of scheduled preventive maintenance strategy of a system

After the joint preventive maintenance time (t^*) for all the elements of the system has been defined, the strategy of their joint maintenance cycle can be easily followed until the failure of one of the elements. When this happens, a decision making problem arises whether the element renewed after the failure should be included in the scheduled joint time (although by then it will not have reached time t^* to the scheduled maintenance), or it should be shifted to the next joint preventive maintenance time (Fig. 1). In the latter case, its time to preventive maintenance would be extended, which would increase the risk of failure.

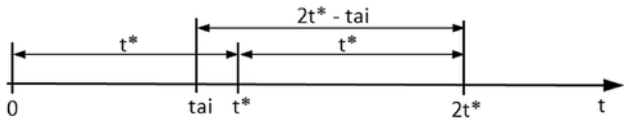


Fig. 1. Element's scheduled time to preventive maintenance and extended after a failure

To solve the problem it is necessary to consider an increased probability of a failure of the element of an extended operation time ($2t^* - t_{ai}$). Moreover, the costs of its potential additional maintenance (k_{oi}) should also be taken into account, when the element's maintenance would take place in the framework of the maintenance of the entire system although by then it will not have reached the scheduled time t^* to its preventive maintenance (it will have reached only $t^* - t_{ai}$). We propose to derive the indicator on the basis of which the decision for the failed element should be taken in such case from the formula:

$$D(t_{ai}) = k_{oi} + F_i\left(t^* - t_{ai}\right) k_{ai} - \frac{F_i\left(2 \cdot t^* - t_{ai}\right) - F_i\left(t^*\right)}{1 - F_i\left(t^*\right)} \cdot k_{ai}\,, \quad (2)$$

where:

- t_{ai} – time of failure of i-th element,
- t^* – joint time to preventive maintenance of elements determined prior to failure of i th element,
- $F_i(t)$ – cumulative distribution function of time to failure of i-th element,
- k_{ai} – cost of failure of i-th element of the system,
- k_{oi} – cost of preventive maintenance of i-th element of the system.

If $D(t_{ai}) \leq 0$, the maintenance of the element together with the other ones is justified, although since the post-failure maintenance at time t_{ai} it has not worked through time t^*. When $D(t_{ai}) > 0$, it is reasonable to extend the working time of the element and exclude it from the nearest maintenance period of the other elements. It should be included in the maintenance procedure during the subsequent maintenance of the system, after the nearest joint maintenance of its elements.

With known failure and maintenance costs and distribution function of times to failure of individual elements, the boundary value of $t_{ai} = t_{i\_gr}$ can be calculated for each of them. It is the time for which $D(t_{ai}) = 0$. If the failure occurs before the element reaches this working time, its additional maintenance together with the other elements in the nearest joint preventive maintenance time is reasonable. If the failure occurs after reaching this working time, it is justified to extend the element's working time with no maintenance. An example of $D(t_{ai})$ plot is shown in figure 2. The curve was plotted following equation (2) for an element of normal distribution of time to failure N(m = 21,5 [months]; σ = 4,75 [months]), with the quotient of failure cost and repair cost ka/ko = 4 and scheduled preventive maintenance time t^*=14 [months].

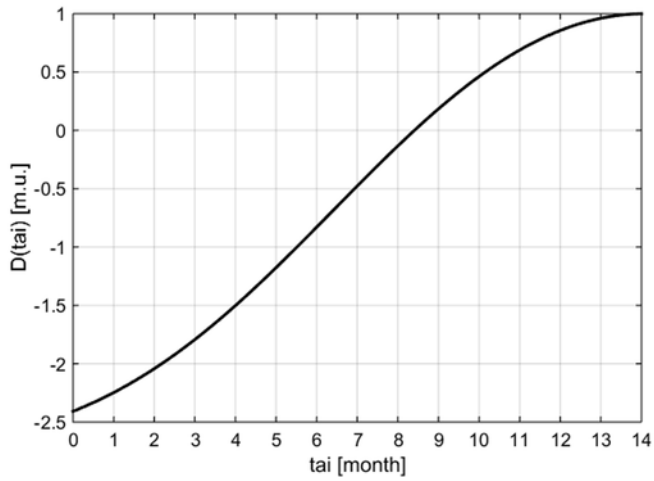


Fig. 2. An example of function $D(t_{ai})$ plot

The value of t_{i_gr} of 8,5 [months] determined for this case is clearly visible. This means that if the element's failure occurs by the 8,5 month of operation, its additional preventive maintenance together with the other elements scheduled in the 14th month is justified. If the failure occurs after the 8,5 month, the period of its operation can be extended (by 5,5 months maximum) and it should be included in the next joint maintenance procedure rather than the nearest one.

Another method of determining t_{i_gr} discussed in the literature [13] is the use of decision-random models based on dynamic programming and Bellman's principle of optimality. With these models preventive maintenance optimal time (t^*) in a finite time horizon (T_h) of the operation of a technical object can be calculated. It is also possible to calculate the highest value of T_h for which the model indicates when preventive maintenance or the system's technical state inspection are not cost effective [13], which is illustrated in the flowchart in figure 3 [18]. The maximum this time (T_{hmax}, with no preventive maintenance or inspection) corresponds to time $2t^* - t_{i\_gr}$ in the model proposed in the present paper. When T_{hmax} is known from the decision-random model also $t_{i\_gr} = 2t^* - T_{hmax}$ can be calculated.

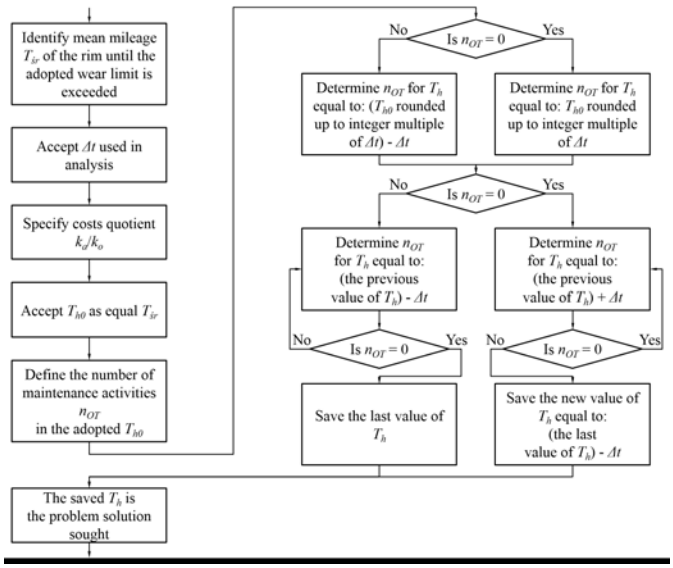


Fig. 3. A method of determining the highest mileage of wheel set rim for which the analysed servicing activities remain uneconomic [18]

What is characteristic of the presented scheduled maintenance strategy is that the initially determined joint preventive maintenance times of all the elements, performed at constant time intervals of t^* remain unchanged. The application of such strategy is particularly justified in cases when the maintenance activities of the given system are performed at constant intervals of time t^* and its multiples connected with other scheduled servicing activities of the entire object or the group of objects.

2.3. A model of adaptive strategy of a system's preventive maintenance

The other proposed maintenance method is an adaptive strategy in which the intervals between the times of subsequent preventive maintenance of the entire system can be modified. Such an interval is re-determined after the failure of any element of the system. In this model the starting situation is the same as in the model of scheduled strategy, that is there is a fixed time to preventive maintenance (t^*) common for all the elements of the system. The strategy compatible with this time is followed until an element fails before reaching age t^*. When this happens, following the adaptive strategy, a post-failure maintenance of the failed component is performed. And if we know that the other elements did not fail by this time (modification of probability distribution of times to failure), another time, common for all the elements, of preventive maintenance is set.

The two maintenance strategies differ from each other and the question which will be more effective in an actual case can be solved by identifying economic indicators resulting from their application in the adopted time horizon [1].

3. A simulation based model of scheduled and adaptive maintenance strategy

The simulation based calculation model was developed on the assumptions formulated for the model of scheduled strategy and the model of adaptive strategy described in section 2. The calculation algorithms enable computer simulations of the operation of a system when each of the strategies is followed. Simplified algorithms showing a single simulation iteration are shown in figures 4a and 4b.

Detailed characteristics of the simulation method used for calculations were presented in [7], where the possibility of estimating the reliability of a complex technical object subjected to decomposition was shown. For the purposes of this paper, this method is a basis that was significantly supplemented with new functions related to preventive renewal of the examined objects. Thanks to this, in addition to the evaluation of the reliability of the object, it is also possible to evalu-

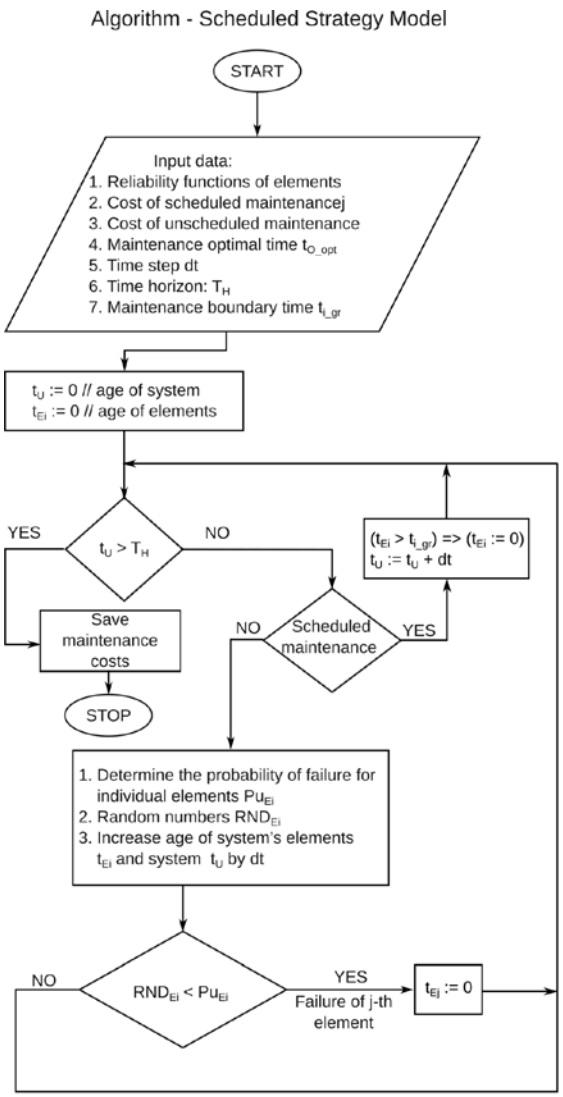


Fig. 4a. Simplified algorithm of a single simulation iteration of system's operation after scheduled strategy model

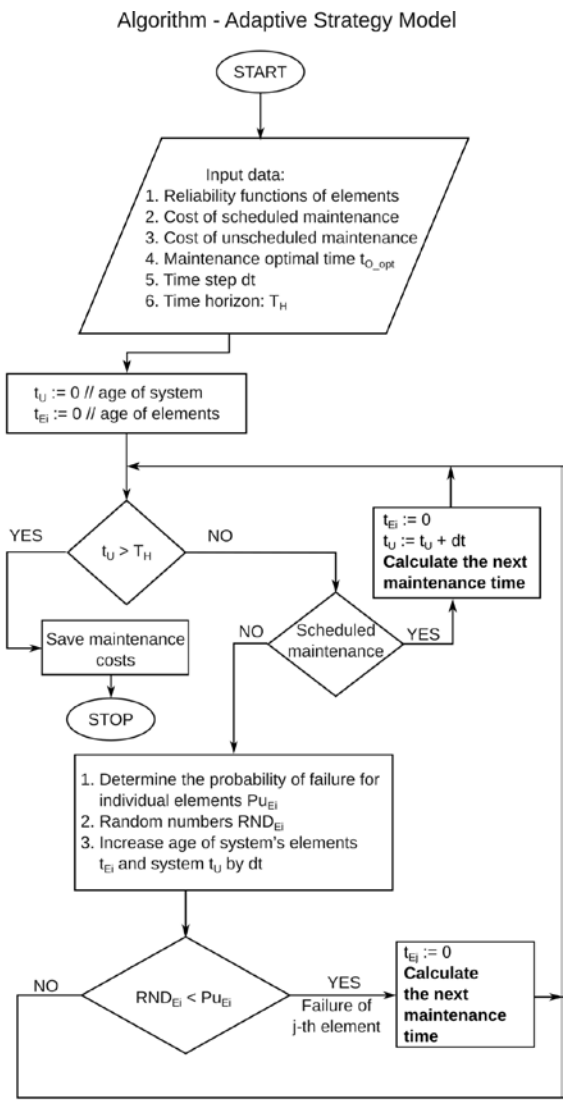


Fig. 4b. Simplified algorithms of a single simulation iteration of system's operation after adaptive strategy model

ate the economic efficiency of various strategies for the preventive renewal of its parts.

The assumptions adopted for the algorithm based on the scheduled strategy model were:

- in the case of a failure of an element, its maintenance is performed with no maintenance of undamaged elements,
- this element will undergo maintenance in the framework of preventive maintenance for time t^*, if its operation time was less than t_{i_gr}.

The input data for the program developed on the basis of this algorithm include: elements' reliability functions, the cost of preventive maintenance of each element k_{oi}, the cost of post-failure maintenance k_{ai}, maintenance optimal time of system's elements t^*, boundary value of time t_{i_gr}, time step dt and time horizon T_H of the simulation and the number of simulations n.

For the algorithm based on the adaptive strategy model the following assumptions were adopted:

- in the case of a failure of an element, its maintenance is performed with no maintenance of undamaged elements and the time of the nearest required preventive maintenance of all elements is calculated,
- preventive maintenance of all elements is performed at a time fixed after each failure of any element (as above).

In this case the only difference in the input data is that the determination of the boundary value of time t_{i_gr} is not necessary.

4. Calculation results

The calculations were conducted on the basis of the data referring to the external research on the monobloc wheel sets of electric rail unit vehicles performed during operational practice.

It was stated that the distribution of the wheel sets' operation time is compatible with Weibull distribution of the shape parameter $\upsilon = 4{,}1$ and scale parameter $\beta = 170000$ [km]. The probability density function of mileage to failure for this distribution is given as equation (3) and presented in figure 5.

$$f(t) = \upsilon\left(\frac{1}{\beta}\right)^{\upsilon} t^{\upsilon-1} exp\left(-\left(\frac{t}{\beta}\right)^{\upsilon}\right), \tag{3}$$

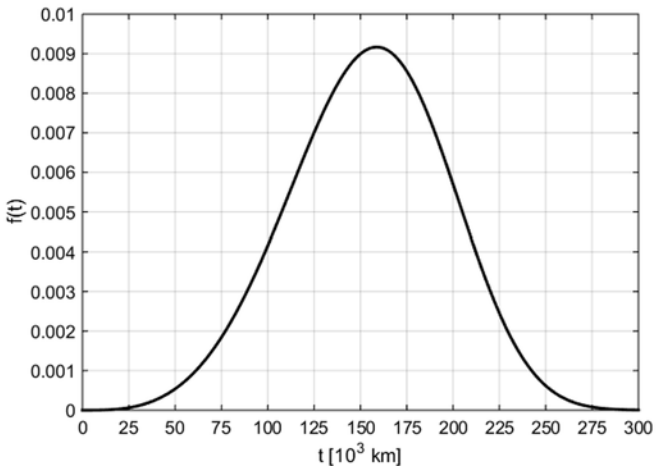


Fig. 5. The probability density function of operation time to failure of the analysed wheel sets

Only failures resulting from wear of vehicle wheels during operation, consisting in exceeding the limit values of wheel profile dimensions specified in the vehicle maintenance system documentation (DSU), are considered. The wheel set operation and maintenance process implemented at the carrier includes inspections and renewals carried out in accordance with DSU. Inspections are carried out at predetermined times - after a specific mileage and time.

Taking into account the costs of: preventive maintenance – k_o = 80000 [m.u. – monetary units] and post-failure maintenance – k_a = 250000 [m.u.] an optimal preventive maintenance time period of the analysed wheel sets was determined. For this purpose the decision-random models were employed [13, 18] and the value of maintenance time was obtained (expressed in mileage kilometres) of $t^* = 108000$ [km].

With these data the curve of function $D(tai)$ was plotted following equation (2), shown in figure 6.

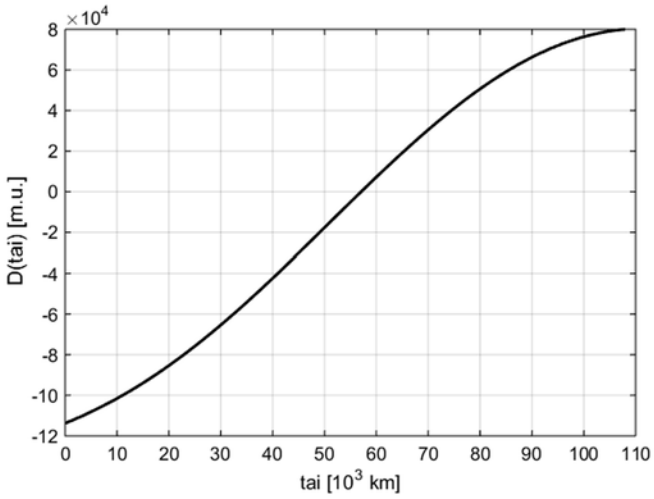


Fig. 6. Function D(tai) for the analysed wheel set

It implies that the boundary value of mileage t_{i_gr} for a wheel set is 58000 [km]. If a failure of a wheel occurs before this mileage is reached, its post-failure maintenance will have to be performed together with the other wheels of the car of the rail unit and again together with all the wheels of the entire system – also in the nearest joint maintenance. If the failure occurs after this time, it is renewed after the failure, but not during the nearest joint maintenance of the wheels of the entire system its operation time is extended until the next maintenance after the nearest joint maintenance.

Using the newly developed simulation models calculations were performed adopting additionally the following values of the parameters for the simulation: $T_H = 1000000$ [km], $dt = 5$ [km], the number of simulation repetitions $n = 10000$.

Some results from the simulation are tabulated in Table 1. The result found the most important was the maintenance costs. This is confirmed in the operational practice in which the economic aspect of operation is frequently the basis for the assessment of the operating system and maintenance strategy adopted [5]. Consequently, the obtained results are significant indications for decision making in the operation process management.

Table 1 shows mean costs of preventive and post-failure maintenance as well as their totalled value, obtained with the application of both strategies of maintenance of the analysed wheel sets in the simulation adopted horizon.

The analysis of the results shows that in the discussed case lower total costs were obtained when the maintenance adaptive strategy was employed. This indicates that its application may be more cost effective in the real life operation of the modelled technical object.

The significant advantages of the use of computer simulation for the analysis of a system's operation following the presented models is the possibility of obtaining the maintenance costs distributions, which is illustrated in figure 7.

Table 1. Results of simulations

Type of value of maintenance costs	Model of scheduled strategy	Model of adaptive strategy
Mean cost of preventive maintenance [m.u.]	614480	618962
Mean cost of post-failure maintenance [m.u.]	1093025	1069500
Mean total cost of maintenance [m.u.]	1707505	1688462

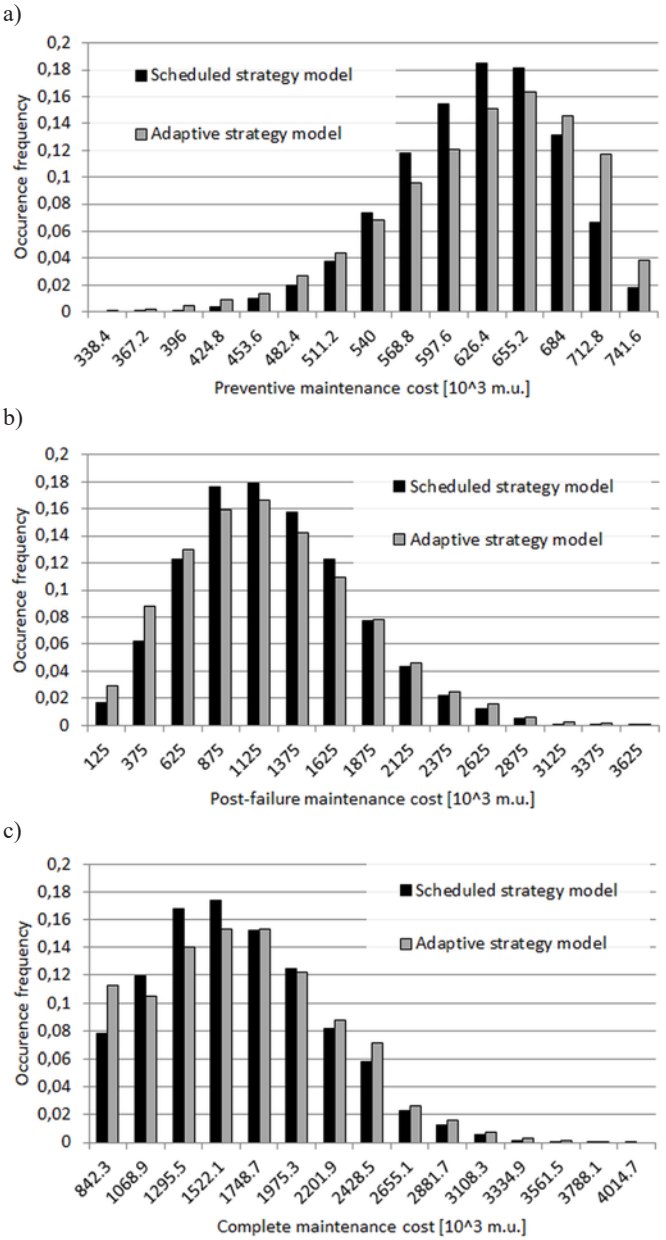


Fig. 7. Distributions of preventive, post-failure and overall maintenance costs for the analysed strategies

Consequently, simulation enables the assessment of the probability of a maintenance cost not exceeding the intended level.

5. Conclusions

The economic indicators of the operation process are one of the most important indicators of the use of vehicles in transport systems. This is because it is the operation management that largely determines the proper functioning of the entire business company. The methods proposed in the paper may significantly affect the rational organisation of inspections and preventive maintenance. Consequently, the choice of a maintenance strategy should include the criterion of cost efficiency. Although the presented strategies do not apparently differ significantly, in actual practice they result in serious economic implications due to considerable costs of maintenance activities.

The presented methods of preventive maintenance scheduling in systems of vehicle operation and maintenance are a useful tool aiding decision making processes and rational use of technical objects. Each model is applicable in various conditions of operation and for complex and structurally differentiated technical objects.

The presented adaptive model of preventive maintenance strategies determination contributes to the development of methods of performing maintenance of complex technical objects.

Research funding

This work was financed by AGH University of Science and Technology, Faculty of Mechanical Engineering and Robotics, research program No. 16.16.130.942.

References

1. Andrzejczak K, Młyńczak M, Selech J. Assessment model of operational effectiveness related to newly operated public means of transport. Proceedings of the 27th European Safety and Reliability Conference ESREL 2017: 3455-3461, https://doi.org/10.1201/9781315210469-435.
2. Badia F G, Berrade M D, Campos C A. Optimal inspection and preventive maintenance of units with revealed and unrevealed failures. Reliability Engineering and System Safety 2002; 78: 157-163, https://doi.org/10.1016/S0951-8320(02)00154-0.
3. Berrade M D, Scarf P A, Cavalcante C A V, Dwight R A. Imperfect inspection and replacement of a system with a defective state: A cost and reliability analysis. Reliability Engineering and System Safety 2013; 120: 80-87, https://doi.org/10.1016/j.ress.2013.02.024.
4. Bradley E. Reliability engineering a Life Cycle Approach. Boca Raton: CRC Press Taylor & Francis Group, 2017, https://doi.org/10.1201/9781315367422.
5. Faulin J, Juan Perez A A, Martorell Alsina S S, Ramirez-Marquez J E (Eds.). Simulation Methods for Reliability and Availability of Complex

Systems. London, New York: Springer, 2010, https://doi.org/10.1007/978-1-84882-213-9.
6. Macchi M, Garetti M, Centrone D, et al. Maintenance management of railway infrastructures based on reliability analysis. Reliability Engineering & System Safety 2012; 104: 71-83, https://doi.org/10.1016/j.ress.2012.03.017.
7. Młynarski S, Pilch R, Smolnik M, Szybka J, Wiązania G. A concept of reliability assessment simulation model using systems structural decomposition. Journal of KONBiN 2018; 46: 51-74, https://doi.org/10.2478/jok-2018-0023.
8. Młynarski S, Pilch R, Smolnik M, Szybka J, Wiązania G. Formation of koon Systems Reliability Estimated with Analytical and Simulation Calculation Methods. Journal of KONBiN 2017; 42: 255-272, https://doi.org/10.1515/jok-2017-0028.
9. Młyńczak M. Failure models of mechanical objects. Zagadnienia Eksploatacji Maszyn 2010; 45: 29-43.
10. Nachlas J A. Reliability engineering. Probabilistic models and maintenance methods. Boca Raton: CRC Press Taylor & Francis Group, 2017, https://doi.org/10.1201/9781315307596.
11. O'Connor P. Practical reliability engineering. Chichester: John Wiley & Sons Ltd., 2012.
12. Peng W, Huang H Z, Zhang X, Liu Y, Li Y. Reliability based optimal preventive maintenance policy of series-parallel systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2009; 2: 4-7.
13. Pilch R. Determination of preventive maintenance time for milling assemblies used in coal mills. Journal of Machine Construction and Maintenance - Problemy Eksploatacji 2017; 1: 81-86.
14. Saranga H, Kumar U D. Optimization of aircraft maintenance/support infrastructure using genetic algorithms ¬level of repair analysis. Annals of Operations Research 2006; 143: 91-106, https://doi.org/10.1007/s10479-006-7374-1.
15. Serkan E, Yilser D. Reliability and optimal replacement policy for a k-out-of-n system subject to shocks. Reliability Engineering & System Safety 2019; 188: 393-397, https://doi.org/10.1016/j.ress.2019.03.045.
16. Song H, Schnieder E. Modeling of railway system maintenance and availability by means of colored Petri nets. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 2: 236-243, https://doi.org/10.17531/ein.2018.2.08.
17. Sowa A. Formal models of generating checkup sets for the technical condition evaluation of compound objects. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(1): 150-157.
18. Smolnik M. Projektowanie procesu obsługiwania obiektów technicznych na przykładzie wybranych wagonów tramwajowych [PhD thesis]. Kraków: AGH w Krakowie, 2018.
19. Świderski A, Jóźwiak A, Jachimowski R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 2: 292-299, https://doi.org/10.17531/ein.2018.2.16.
20. Werbińska-Wojciechowska S. Preventive Maintenance Models for Technical Systems. In: Technical System Maintenance. Springer Series in Reliability Engineering. Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-10788-8.
21. Zhao Y X. On preventive maintenance policy of a critical reliability level for system subject to degradation. Reliability Engineering and System Safety 2003; 79: 301-308, https://doi.org/10.1016/S0951-8320(02)00201-6.

Stanisław MŁYNARSKI
Cracow University of Technology
Institute of Rail Vehicles
Jana Pawla II 37, 31-864 Cracow, Poland

Robert PILCH
Maksymilian SMOLNIK
Jan SZYBKA
Grzegorz WIĄZANIA
AGH University of Science and Technology
Faculty of Mechanical Engineering and Robotics
Al. A. Mickiewicza 30, 30-059 Cracow, Poland

Emails: mlynarski_st@poczta.onet.pl, pilch@agh.edu.pl, smolnik@agh.edu.pl, szybja@agh.edu.pl, wiazania@agh.edu.pl

Article citation info:
CUI J, REN Y, XU B, YANG D, ZENG S. Reliability analysis of a multi-ESO based control strategy for level adjustment control system of quadruped robot under disturbances and failures. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 42–51, http://dx.doi.org/10.17531/ein.2020.1.6.

Jingjing CUI
Yi REN
Binghui XU
Dezhen YANG
Shengkui ZENG

RELIABILITY ANALYSIS OF A MULTI-ESO BASED CONTROL STRATEGY FOR LEVEL ADJUSTMENT CONTROL SYSTEM OF QUADRUPED ROBOT UNDER DISTURBANCES AND FAILURES

ANALIZA NIEZAWODNOŚCI STRATEGII STEROWANIA W OPARCIU O KILKA OBSERWATORÓW STANU ROZSZERZONEGO ESO I JEJ ZASTOSOWANIE W SYSTEMIE KONTROLI REGULACJI POZIOMU CZWORONOŻNEGO ROBOTA W WARUNKACH ZAKŁÓCEŃ I USZKODZEŃ

The complexity of control algorithms and their vulnerability to disturbances and failures are the main problems that restrict the operations of multi-legged mobile robots in more complex environments. In this paper, a multiple extended state observer (ESO) based control strategy is proposed to achieve stable tilt angle control for quadruped robots under the influence of disturbances and actuator failures. By treating the multiple legs as parallel control objects, more ESOs were added to improve the disturbance rejection ability of the linear active disturbance rejection control (LADRC). Correlation of interactive information about the legs is realized by the synthesis of multiple ESO information. Based on LADRC, this method has the advantages of easy parameter tuning, good robustness, and strong ability to cope with interference and fault conditions. A control system reliability evaluation method was proposed. The reliability and control performance of the multi-ESO based control system under leg stuck failure conditions were systematically analyzed. Simulation and experimental results for the level adjustment control system of a quadruped robot are provided to verify the disturbance rejection ability, feasibility and practicability of the proposed multi-ESO based control method.

***Keywords**: reliability analysis, fault tolerance, active disturbance rejection control (ADRC), quadruped robots.*

Złożoność algorytmów sterowania oraz ich brak odporności na zakłócenia i uszkodzenia to główne czynniki ograniczające pracę wielonożnych robotów mobilnych w bardziej złożonych środowiskach. W przedstawionym artykule zaproponowano strategię sterowania wykorzystującą kilka obserwatorów stanu rozszerzonego (ESO), która pozwala na uzyskanie stabilnego kąta pochylenia robota czworonożnego w warunkach zakłóceń i uszkodzeń siłowników. Traktując każdą z nóg jako równorzędny obiekt sterowania, dodano dodatkowe ESO, co pozwoliło na poprawienie zdolności algorytmu liniowego aktywnego tłumienia zakłóceń (LADRC) do kompensacji (tłumienia) tych ostatnich. Interaktywne informacje dotyczące poszczególnych nóg korelowano poprzez syntezę danych z ESO. Zaletami omawianej metody opartej na LADRC są: łatwość dostrajania parametrów, wysoka niezawodność oraz bardzo dobra zdolność do radzenia sobie z zakłóceniami i uszkodzeniami. Zaproponowano także metodę oceny niezawodności systemu sterowania. Analizowano niezawodność i wydajność systemu opartego na kilku ESO w warunkach awarii wywołanej zablokowaniem nóg robota. Przedstawiono wyniki badań symulacyjnych i eksperymentalnych systemu sterowania regulacją poziomu robota czworonożnego, które pozwalają zweryfikować zdolność proponowanej metody do tłumienia zakłóceń, a także możliwość jej praktycznego zastosowania.

***Słowa kluczowe**: analiza niezawodności, tolerancja uszkodzeń, aktywne tłumienie uszkodzeń (ADRC), roboty czworonożne.*

1. Introduction

With the rapid development of technology and increase in performance requirements, the complexity of multi-legged robots has increased in terms of mechanical structures and control algorithms, which has led to challenges in guaranteeing the reliability of the systems. To achieve reliability requirements, stable control under the influence of disturbances and failures is critical for multi-legged mobile robot systems. Researchers have made efforts to promote the reliability and disturbance rejection ability of mobile robot control systems. Intelligent distributed control [3,17,22,34] and fault-tolerant control [8,23-25] are the main solutions. Christensen et al. [2] designed a distributed control strategy for self-reconfigurable modular robots. Cully et al. [5] proposed an intelligent self-adaption control method for a six-legged robot. The researchers attempted to use intelligent distribution methods to address interactions and information crossover between multiple actuators. In the literature [4], a series control scheme based on active disturbance rejection control (ADRC)

is presented for mobile robots to improve their reliability under actuator failure without changing parameters, but only a 2-dimensional tilt angle of the mobile robot is considered. Through mathematical formula derivations, Kommuri [12] designed an observer-based sensor fault-tolerant control for electric vehicles. Xiao and Yin [27] designed a sliding mode observer to enhance fault-tolerant control of nonlinear systems with disturbance and actuator faults. Observer-based control methods can effectively capture the disturbances and faults caused by an uncertain environment, and then, the corresponding fault-tolerant control algorithms can be designed.

Although these methods have a certain effect on improving system reliability, the design process needs to obtain enough information, including object functions, mechanical dynamics, possible failure modes, effects, reasons and modeling of failures. Complicated analysis and deducing processes are also required. Reliability analysis and design for non-monotonic and dynamic control systems is difficult due to the complexity of the control algorithms. Simple methods are still needed due to the problems of uncertainty interference and failures. In addition, reliability analysis of control systems is still focused on FMECA or reliability tests for key objects or structures [19,32]. The reliability of dynamic control systems lacks quantitative evaluation methods.

The extended state observer (ESO) was first proposed by Han in [9] and has been the basis to form a new control technology that is known as active disturbance rejection control (ADRC). The total disturbance, including internal nonlinear dynamics, coupling effects, model inaccuracy, and external disturbance and uncertainties, is regarded as an extended state of the system to be observed. The estimated disturbance obtained by ESO will be used to compensate the control output to achieve disturbance rejection. Control of large uncertainties together with satisfactory performance can be achieved by ADRC due to the strong capability of ESO [29].

Since ESO was proposed, Gao [13] and some other scholars have made many efforts to improve ESO and ESO based control strategies. Fractional ADRC [13], linear ESO and linear ADRC [6], linear-nonlinear switching ADRC [15] and some other control methods based on ESO [16,20] were proposed. Yang et al. [31] used enhanced ESO and corresponding ADRC for a MIMO system, which could accelerate the response time. Xue et al. [29] proposed an adaptive ESO and applied it to air–fuel ratio control in gasoline engines. In [7] and [28], ESO was designed and utilized for fault diagnosis and fault-tolerant control. Thus far, applications of ESO based control strategies have proven its superiority in dealing with uncertainties and disturbances [1,14,18]. Moreover, the stability of ESO has also been analyzed and theoretically proven [11]. Performance [21,30] and limits [10,26] were also analyzed. Considering the merits of ESO, the distributed ESO based control strategy provides an effective solution to the uncertain disturbance and failure problems of the multi-legged mobile robot level adjustment control system.

In this paper, a multi-ESO based control strategy is proposed to promote the reliability of the quadruped robot level adjustment control system during disturbances and actuator failures. For parallel control objects, such as multiple legs, more ESOs were added to improve the disturbance rejection ability of the LADRC. Based on multiple linear ESOs, this novel method has superior parameter tuning and ability to deal with uncertainties. The information interaction between the actuators can be achieved by the integration of information from multiple ESOs. Motion reliability indicators have also been proposed to analyze and evaluate the reliability of motion control systems. A quantitative reliability evaluation method for the dynamic response of the control systems is designed and proposed. Simulation and experimental results show the satisfactory performance and reliability of the proposed method for addressing disturbances and failures.

The rest of this paper is organized as follows. Section 2 describes the design and process of the multi-ESO based control strategy in detail. Section 3 gives the control system reliability evaluation method. Section 4 introduces the plant, states the modeling process, and classifies typical failure conditions. Section 5 presents the simulation results and reliability analysis systematically. Conclusions about the presented technique are given in Section 6.

2. Multi-ESO Based Control Method

2.1. Linear Active Disturbance Rejection Control

Nonlinear functions were designed and used in the original ADRC controller proposed by Han [9]. Although the performance of the nonlinear ADRC controller is very good, the parameter adjustment is complicated, and it is difficult to achieve simple and rapid control in practical applications. Gao [13 15] found through a large number of studies that proper linear functions can also obtain excellent performance for ADRC controllers, and the calculation of parameter tuning is greatly reduced, which is more suitable for engineering applications. The basic idea of the ADRC linear simplification is to linearize the ESO and relate its parameters to the observer bandwidth to simplify the ESO design. A simple PD control is used to correlate the proportional coefficient and the derivative time constant with the controller bandwidth to simplify controller tuning. A standard single ESO LADRC control scheme is shown in Fig. 1.

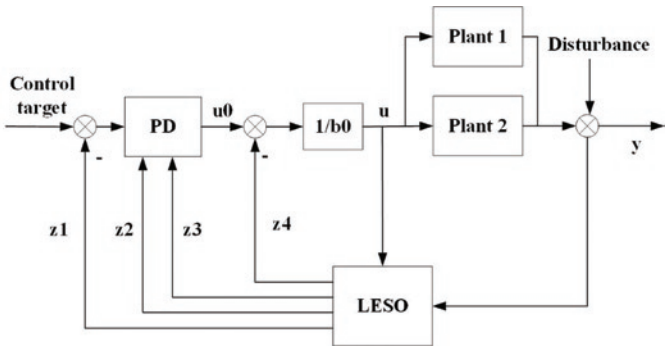


Fig. 1 Standard single ESO LADRC control scheme.

2.2. Design of the Extended State Observer

The linear extended state observer (LESO) is based on a simplified linear structure and uses the input and output of a system to estimate the extended system states. Because ADRC has the characteristics of being independent of the object model, the LESO can be designed as standard object model-free forms. In the case where a partial model of the object is known, a model-assisted LESO can be designed based on the known model information. Both the calculation of the ESO and the lag of the disturbance estimate can be reduced. Moreover, for an n-order system, the states estimated by the ESO contain the disturbance information, the system output and the 1 to n-1 order derivatives. When the system output and some low-order derivatives can be obtained, it is unnecessary to use the ESO to estimate these state variables; thus, the LESO structure can be simplified to obtain the reduced-order LESO. Standard object model-free LESO is adopted in this paper.

For an n-order system:

$$y^{(n)} = -a_{n-1}y^{(n-1)} - a_{n-2}y^{(n-2)} - \cdots - a_0 y + \omega + bu \tag{1}$$

where u is the input, y is the output, and ω is the disturbance. a_0, $a_1, \ldots a_{n-1}$ are object structure parameters, which are unknown here. b is part known and the known part is denoted as b_0. Then, the above formula can be written as:

$$y^{(n)} = f + b_0 u \tag{2}$$

where f is the total disturbance of the system, including external and internal disturbances.

By expanding the total disturbance f to the n+1 state variable of the system, equation (2) can be transformed into an n+1 order extended state space description:

$$\begin{cases} \dot{x} = Ax + Bu + E\dot{f} \\ y = Cx \end{cases} \tag{3}$$

In which,

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}_{(n+1)\times(n+1)} \quad B = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ b_0 \\ 0 \end{bmatrix}_{(n+1)\times 1} \quad E = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \\ 1 \end{bmatrix}_{(n+1)\times 1}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}_{1\times(n+1)}$$

The corresponding LESO is:

$$\begin{cases} \dot{z} = Az + Bu + L(y - Cz) \\ \dot{y} = Cz \end{cases} \tag{4}$$

where $z \to x$ is the observer state vector and L is the observer error feedback gain matrix. Using $L = \begin{bmatrix} l_1 & l_2 & \cdots & l_{n+1} \end{bmatrix}^T$, the observer characteristic polynomial is:

$$s^{n+1} + l_1 s^n + \cdots + l_{n+1} = (s + \omega_0)^{n+1} \tag{5}$$

Discretization is needed for computer simulation. The object of this paper is a third-order system. The Euler method is used to discretize the LESO, and the fourth-order discrete LESO is constructed as:

$$\begin{cases} z(k+1) = \Phi_E z(k) + \Gamma_E u_d(k) \\ y_d(k) = \mathrm{H}_E z(k) + L_E u_d(k) \end{cases} \tag{6}$$

where:

$$\Phi_E = \begin{bmatrix} 1-(l_1+l_2h) & h & 0 & 0 \\ -(l_2+l_3h) & 1 & h & 0 \\ -(l_3+l_4h) & 0 & 1 & h \\ -l_4 & 0 & 0 & 1 \end{bmatrix} \quad \Gamma_E = \begin{bmatrix} 0 & l_1+l_2h \\ 0 & l_2+l_3h \\ b_0h & l_3+l_4h \\ 0 & l_4 \end{bmatrix}$$

$$\mathrm{H}_E = \begin{bmatrix} 1-l_1h & 0 & 0 & 0 \\ -l_2h & 1 & 0 & 0 \\ -l_3h & 0 & 1 & 0 \\ -l_4h & 0 & 0 & 1 \end{bmatrix} \quad L_E = \begin{bmatrix} 0 & l_1 \\ 0 & l_2 \\ 0 & l_3 \\ 0 & l_4 \end{bmatrix} \quad u_d(k) = \begin{bmatrix} u(k) \\ y(k) \end{bmatrix}$$

And h is the system sampling time.

Let $\beta = e^{-\omega_0 h}$, and then:

$$\begin{cases} l_1 = 1 - \beta^4 \\ l_2 = \frac{1}{h}(\beta^4 - 4\beta + 3) \\ l_3 = \frac{1}{h^2}(-\beta^4 + 6\beta^2 - 8\beta + 3) \\ l_4 = \frac{1}{h^3}(4\beta^4 - 4\beta^3 + 5\beta^2 - 4\beta + 1) \end{cases} \tag{7}$$

2.3. Design of the Multi-ESO Based Control Strategy

In this section, a multi-ESO based linear active disturbance rejection control strategy is proposed. The anti-interference ability of the LADRC is utilized, and the system can cope with the overall environmental disturbance. However, the control signals sent to a group of plants are fixed and cannot be adjusted according to the situation of each plant, so the performance under actuator failure may be unsatisfactory. The control strategy we proposed designs an ESO for each plant, and the PD part shares the same. Considering the same structure of plants in this paper, the parameters of each ESO can be the same as well, which will not increase the number of parameters of the entire LADRC control strategy. The control signals received by each plant can be dynamically adjusted based on different state observations and estimates, which may further improve the disturbance rejection and fault tolerance of the LADRC. The structure of the proposed multi-ESO based LADRC strategy is shown in Fig. 2.

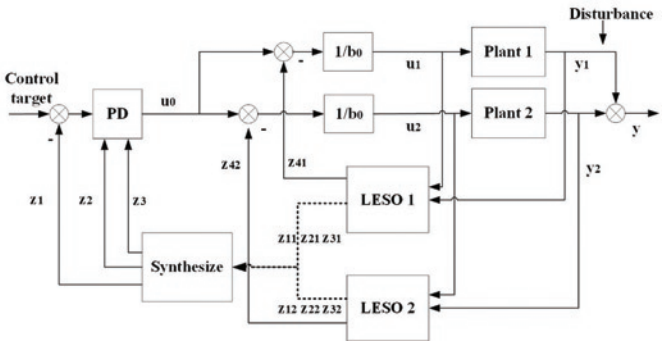


Fig. 2 The multi-ESO based LADRC strategy.

In the multi-ESO based control strategy, LESO 1 and LESO 2 use the design in Section 2.2. Here, the input and output of LESO 1 are u_1, y_1 and z_{11}, z_{21}, z_{31}, z_{41}, respectively. The input and output of LESO 2 are u_2, y_2 and z_{12}, z_{22}, z_{32}, z_{42}, respectively. As the estimated total disturbances of plant 1 and plant 2, z_{41} and z_{42} are used to compensate the control signals. z_1, z_2, and z_3 are the system states estimated by the observers after synthesis. Since $z_{11} \to y_1, z_{21} \to \dot{y}_1, z_{31} \to \ddot{y}_1, z_{12} \to y_2, z_{22} \to \dot{y}_2, z_{32} \to \ddot{y}_2$ and $z_1 \to y, z_2 \to \dot{y}, z_3 \to \ddot{y}$, the synthesized part should be designed according to the relationship among y_1, y_2, and y. In this paper, $y = y_1 + y_2$, therefore:

$$\begin{cases} y = y_1 + y_2 \\ \dot{y} = \dot{y}_1 + \dot{y}_2 \\ \ddot{y} = \ddot{y}_1 + \ddot{y}_2 \end{cases} \tag{8}$$

and

$$\begin{cases} z_1 = z_{11} + z_{12} \\ z_2 = z_{21} + z_{22} \\ z_3 = z_{31} + z_{32} \end{cases} \tag{9}$$

The PD part is designed as follows:

$$u_0 = k_p(r - z_1) - k_{d1}z_2 - k_{d2}z_3 \tag{10}$$

where r is the control target. The controller gain matrix is $\mathbf{K}=[k_{d2} \quad k_{d1} \quad k_p]^T$, and thus, the system characteristic polynomial is:

$$s^3 + k_{d2}s^2 + k_{d1}s + k_p = (s + \omega_c)^3 \tag{11}$$

After discretization, the output u_0 of the PD part is:

$$u_0 = \omega_c{}^3 z_1 - 3\omega_c{}^2 z_2 - 3\omega_c z_3 \tag{12}$$

Then, the control signal to the actuators is:

$$u_1 = \frac{u_0 - z_{41}}{b_0} \tag{13}$$

$$u_2 = \frac{u_0 - z_{42}}{b_0} \tag{14}$$

Thus, the multi-ESO based control strategy only needs to adjust the controller bandwidth ω_c, the observer bandwidth ω_0, and b_0 for a total of three parameters. The controller bandwidth ω_c determines the response speed of the controller. When the value is within an appropriate range, the larger the value of ω_c, the better the system control effect is. However, excessively large ω_c may cause the system to be unstable. The observer bandwidth ω_0 affects the tracking speed of the state observer. The tracking speed of the state observer increases with increasing ω_0, but an overly large speed could cause oscillation or noise. b_0 represents the characteristics of the object.

3. Control System Reliability Evaluation Method

To analyze and evaluate the reliability of the control systems after failure occurs, the following reliability evaluation indicators are given:

(1) Stability reliability, R_s

$$R_s = \frac{U_M}{U_R} \tag{15}$$

In which U_M is the stable controlled angular range under failures, U_R is the demanded stable controlled angular range.

The stability reliability reflects the controllability of the control system under fault conditions. Within the stable controlled angular range, the system is still stable and will not have a fatal stable fault.

(2) Steady state reliability, R_r

$$R_r = \begin{cases} \frac{1-e_f}{1-e_o} & (e_f > e_o) \\ 1 & (e_f \le e_o) \end{cases} \tag{16}$$

where e_o is the steady-state error of the control system without faults, and e_f is the steady-state error of the control system under fault conditions (with the maximum disturbance under control).

Steady-state reliability reflects the ability of the control system to maintain control accuracy under fault conditions.

(3) Transient reliability, R_t

$$R_t = \begin{cases} \frac{1-\sigma_f}{1-\sigma_o} & (\sigma_f > \sigma_o) \\ 1 & (\sigma_f \le \sigma_o) \end{cases} \tag{17}$$

In which σ_o is the overshoot of the control system without faults, and σ_f is the overshoot of the control system under fault conditions (with the maximum disturbance under control).

Transient reliability reflects the stability of the control system under fault conditions.

(4) Time coefficient, t_R

$$t_R = \begin{cases} \frac{t_{so}}{t_{sf}} & (t_{sf} > t_{so}) \\ 1 & (t_{sf} \le t_{so}) \end{cases} \tag{18}$$

where t_{so} is the response time of the control system without faults, t_{sf} is the response time of the control system under fault conditions (with the maximum disturbance under control).

The time coefficient reflects the ability of the control system to maintain the response speed under fault conditions.

Based on the analysis above, the concept of the motion reliability of the motion control system can be proposed to evaluate the reliability of the control system.

Definition: The motion reliability of a motion control system refers to the ability of a system to maintain its control performance under specified conditions when a fault occurs, recorded as R_m.

This is calculated by:

$$R_m = R_s * R_r * R_t * t_R \tag{19}$$

4. System Analysis and Plant Modeling

4.1. Level Adjustment Control System of a Quadruped Robot

Mobile robots require good obstacle crossing ability and reliability to adapt to different terrains. A quadruped robot design is adopted and analyzed in this paper. A quadruped robot is equipped with four legs that can be moved up and down at the four corners of a square platform. The two legs on the diagonal are a group that can change the level adjustment of the robot platform in the diagonal direction. The main parts of the quadruped robot level adjustment control system are shown in Fig.3. In the level adjustment control system, the pose signal of the robot is detected by an altitude angle sensor and sent to the computer. Then, the control signal is sent to the controller after calculation. The legs follow the control commands to make the corresponding movements under the motor drive. When a platform slant is detected, the legs deform and attempt to reduce the tilt angle. This design helps the robot reduce the possibility of rollover and enhances stability when encountering obstacles.

The robot platform performs 3-dimensional space motion on the ground, and its degree of freedom is 3. To describe the altitude change of the platform, a space coordinate system needs to be established.

We use the following assumptions:

(1) The robot platform is a rigid body;
(2) The geometric center and center of gravity of the platform are coincident;
(3) The change in the heading angle of the robot platform during movement is small and can be ignored.

Based on the above assumptions, the established spatial coordinate system describing the altitude of the platform includes two coor-

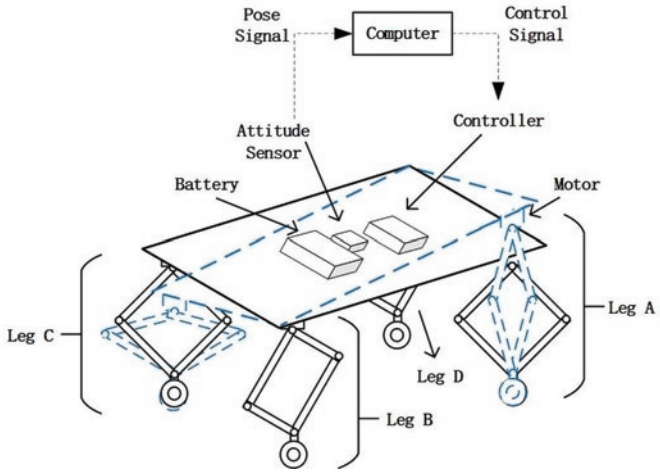


Fig. 3. Main parts of a quadruped robot level adjustment control system

dinate systems, the inertial coordinate system E-XYZ and the follower coordinate system B-XYZ, which are shown in Fig. 4. The origin of E-XYZ is the geometric center of the platform when the height of the platform is half of the maximum height from the ground. East of the origin is the X-axis positive direction (*Xe*), north is the Y-axis positive direction (*Ye*), and vertical to the ground is the Z-axis positive direction (*Ze*). For the B-XYZ coordinate system, the origin is at the geometric center of the platform. The direction of the platform geometric center pointing to leg A is the X-axis positive direction (*Xb*), the direction of the geometric center pointing to leg B is the Y-axis positive direction (*Yb*), and the direction perpendicular upward to the plane formed by *Xb* and *Yb* and through the geometric center is the Z-axis positive direction (*Zb*).

Define $q=(\Phi, \Theta, \Psi, x, y, z)^T$ as the generalized coordinate vector of the robot platform relative to the inertial coordinate system E-XYZ. $\eta=(\Phi, \Theta, \Psi)^T$ is the Euler angle that describes the attitude angle, in which Φ, Θ, Ψ represent the roll angle, pitch angle, and heading angle, respectively. $X=(x, y, z)^T$ stands for the spatial position vector of the platform centroid relative to the inertial coordinate system.

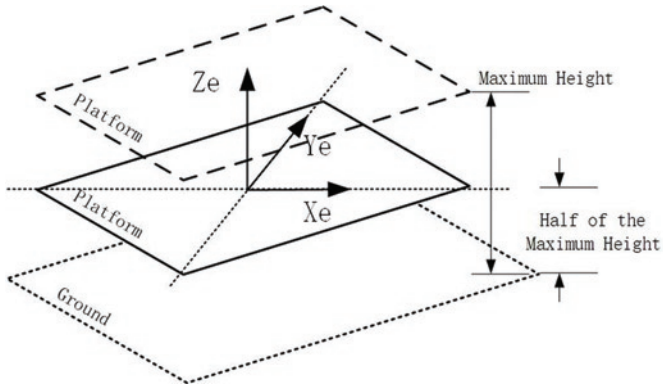
The length of legs A and C of the robot will control the roll angle of the platform. When the length of leg A is greater than that of leg C, the platform will have a positive roll angle ($\Phi>0$). In contrast, if the length of leg A is less than that of leg C, the platform will have a negative pitch angle ($\Phi<0$).

Similarly, the length of legs B and D on the robot will control the pitch angle of the platform. When the length of leg B is greater than that of leg D, the pitch angle is $\Theta>0$. When the length of leg B is less than that of leg D, the pitch angle is $\Phi<0$.

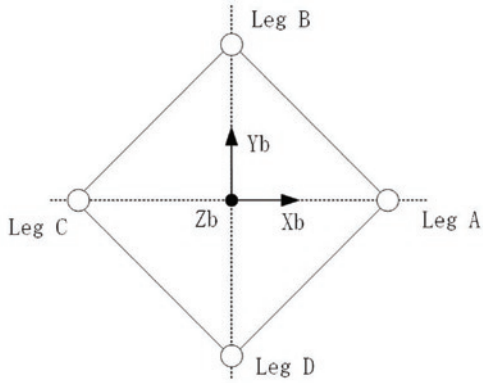
4.2. Plant Modeling

Each leg of the robot consists of two identical branches, each of which includes the upper part and the lower part. The structure of a leg is shown in Fig. 5. The length of the upper part is *L1,* and the lower part is *L2*. The two upper parts of each leg are driven by the motors and rotate around the *O* point for active motion. The two lower sections are hinged together for passive motion. Theoretically, the angle change of ϕ_1 and ϕ_2, which is the change in the rotation angle of the two upper parts, can realize the length change between *L2-L1* and *L2+L1* for each leg.

The whole structure of the leg turns the electric energy of a battery into the rotational angular velocity of the leg. According to the bond graph theory, the bond graph model for each leg can be obtained as shown in Fig. 6. Bond graph models can be transformed into a block diagram and simulated [33]. Using this method, a block diagram simulation model of the leg structure can be obtained without deriving its transfer function. The dynamic constraints of the object are embedded



a) The E-XYZ coordinate system



(b) The B-XYZ coordinate system

Fig. 4. Coordinate systems

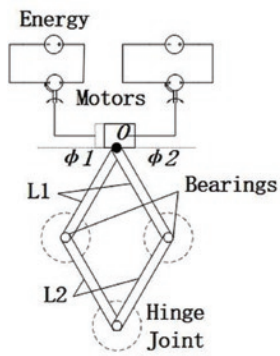


Fig. 5. The structure of a leg

Table1. List of Model Parameters

Description	Symbol	Value
Motor input voltage	Se	12 V
Motor output gear radius	r	0.025 m
Motor inductive reactance	R	1 Ω
Bearing and Hinge damping	b1 b2	4 N.s/m
Leg elastic coefficient	k	8 N/m
Length of the upper part	L1	0.1 m
Mass of the upper part	m1	0.1 kg
Length of the lower part	L2	0.1 m
Mass of the lower part	m2	0.1 kg
Bearing mass	m	0.01 kg

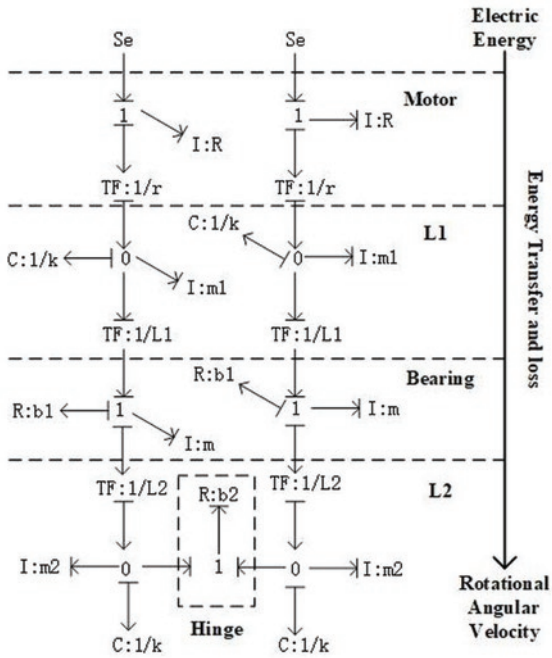


Fig. 6. Bond graph model of a leg

in the bond graph modeling rules. The parameters used in the simulation model are shown in TABLE 1.

4.3. Classification of Failure Conditions

Legs or motors being stuck is the most likely failure mode during operation of the level adjustment control process of a quadruped robot. Leg stuck failure will lead to a lower ability to resist disturbances or the inability to pass obstacles. We determined three kinds of leg stuck failure states in the control process of the quadruped robot, denoted by states A, B, and C.

Failure state A indicates that the failure occurs at the same time as the disturbance. The leg will be at the initial position and cannot follow the instructions of the controller.

Failure state B indicates that stuck failure occurs when the leg moves to the maximum length and stops during the control process.

Failure state C indicates that the failure occurs after one interference and when facing another. The leg will be stuck at the last position after one adjustment. Then, a disturbance is applied again to the control system. This failure state includes two cases, i.e., dealing with a small interference first and dealing with a large interference first.

The failure conditions during the level adjustment control process are shown in Fig.7. The latter simulation analysis will analyze the

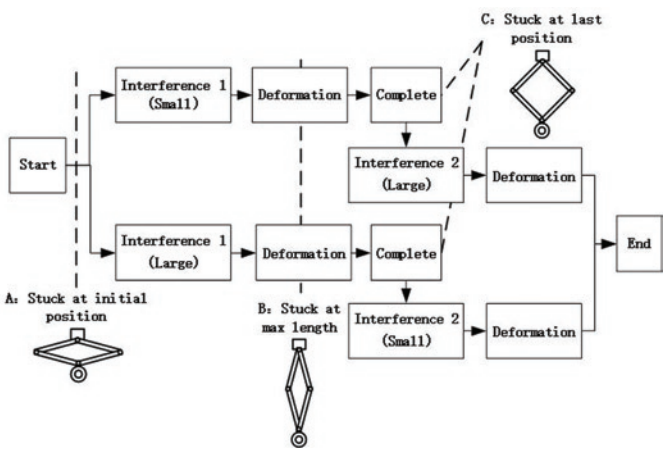


Fig.7 Failure conditions during the level adjustment control process

control performance and reliability of the proposed method from the perspective of these three types of faults.

5. Reliability Analysis

5.1. Simulation Results

With the system model and the multi-ESO based control strategy, the simulation was established in MATLAB/simulink. Considering the ability of the actual system and the actuators, the corresponding limiting module was added to the system simulation model. With the strong robustness and performance of the ESO based control method, the parameters were easily obtained through tuning in the simulation, and relatively good parameters were set as ω_c=21, ω_0=42, and b_0=380. The standard LADRC control scheme uses the same parameters as the multi-ESO-based control method.

Based on the information above, the simulation curves under normal conditions without leg failures are plotted in Fig. 8. The dashed lines in the figure are the responses of the standard LADRC control scheme, and the solid lines are the responses of the proposed multi-ESO based control strategy. The results illustrate that the two control strategies perform well without leg failures under disturbances of 10° and 20°. The response time is less than 2 seconds. Overall, the responses of the new method are slightly inferior to those of the standard LADRC due to the increased complexity of the structure in the multi-ESO based control method. For failure-free situations that do not require flexible changes of the control scheme, the simple method performs better. Under the 20° disturbance, the e_o of the LADRC and multi-ESO control is 0, the σ_o of LADRC is 12.5% and the σ_o of multi-ESO is 15%, and the t_{so} of LADRC and multi-ESO is 1.5 s. Next, the performance of the control strategies under leg stuck failure conditions will be analyzed and discussed.

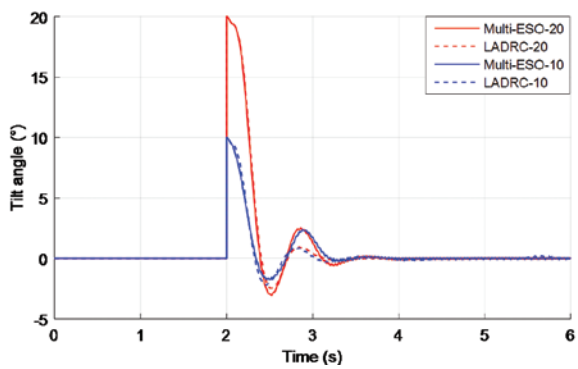


Fig. 8 Simulation results under normal conditions.

(1) Failure state A

In failure state A, the leg is stuck in the initial position and cannot follow the instructions of the controller, which may reduce the disturbance rejection ability of the level adjustment control system. The performances of the multi-ESO based control strategy and the standard LADRC scheme in failure state A are shown in Fig. 9. Both methods can handle 20 degree tilt angle interference under the conditions of failure state A. The response time of the systems when failure state A occur appears to be longer than that under normal conditions, which takes nearly 3 seconds. Under small angle interference (10°), the performances of the two control methods are basically the same. However, when the interference is large (20°), the response of the multi-ESO based control method is significantly better than that of the standard LADRC, with a smaller amplitude and faster response time. In view of the performances of the two methods, it is better to adopt the proposed multi-ESO based control strategy in failure state A. In this situation under 20° disturbance, the e_f of LADRC and multi-ESO

control is 0, the σ_f of LADRC is 32.5% and σ_f of multi-ESO is 12.5%, and the t_{sf} of LADRC is 2.5 s and that of multi-ESO is 1.8 s.

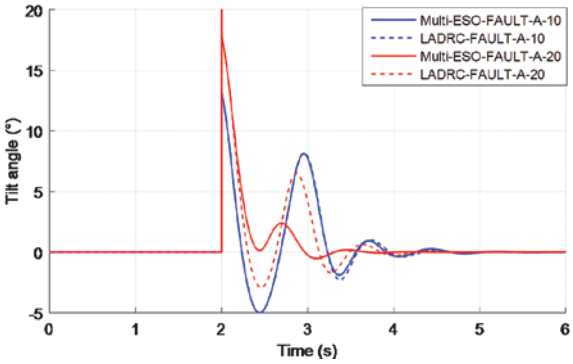
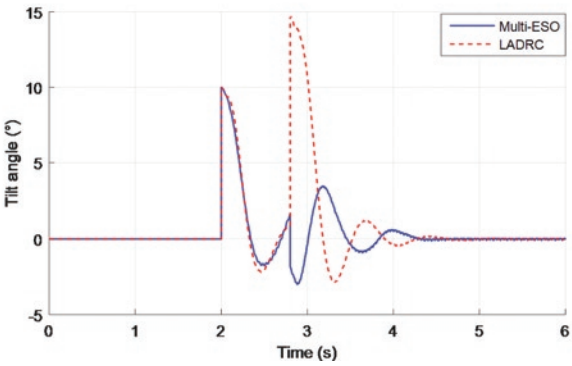


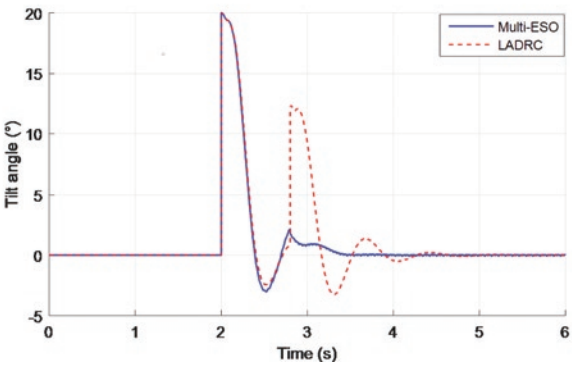
Fig. 9. Simulation results in failure state A

(2) Failure state B

In failure state B, the leg is stuck at the maximum length during the control process. This kind of large angle stuck failure may cause a significant drop in the system disturbance rejection ability. Fig. 10 shows the system responses of the multi-ESO based control strategy and the standard LADRC control scheme in failure state B. Both methods can handle 20 degree tilt angle interference under failure state B, and the response time is a little longer than that under normal conditions. The oscillation of the system response is much worse than that under normal conditions. The responses to the 20 degree interference are better than those to the 10 degree interference because the stuck position is close to the required position of the controller to the leg to resist the disturbance. However, the anti-interference ability of the system still decreases under failure state B. It can be seen clearly that the control quality of the proposed multi-ESO based control strategy is quite good and has advantages over the standard LADRC when failure state B occurs. In this situation, under a 20° disturbance, the e_f of LADRC and multi-ESO control is 0, the σ_f of LADRC is 60% and the σ_f of multi-ESO is 15%, and the t_{sf} of LADRC is 2.6 s and that of multi-ESO is 1.6 s.



(a) Interference of 10 degrees

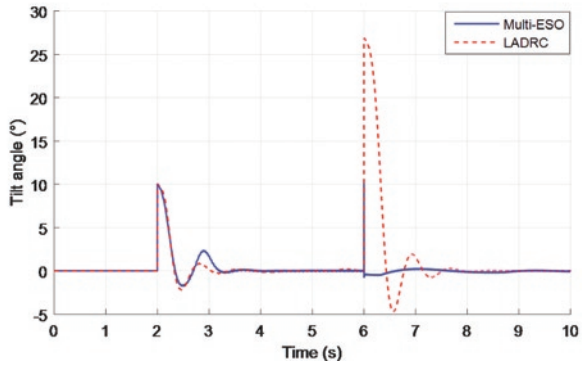


(b) Interference of 20 degrees

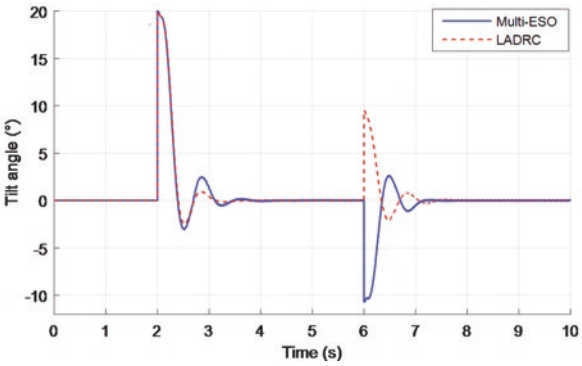
Fig. 10. Simulation results in failure state B

(3) Failure state C

In failure state C, the leg is stuck at the last position after one adjustment, and then, an interference is applied again to the control system. The simulation curves of the multi-ESO based control strategy and the standard LADRC scheme in failure state C are shown in Fig. 11. In Fig. 11(a), a small interference of 10 degrees was applied to the systems first, followed by a large interference of 20 degrees. In Fig. 11(b), the systems are dealing with the large interference first. Both methods are capable of handling 20 degree interference in both cases of failure state C. It is obvious that the control performance of the proposed multi-ESO based control is significantly stronger than that of the standard LADRC in the case of dealing with a small disturbance first and then a large disturbance. In this situation, the e_f of LADRC and multi-ESO control is 0, the σ_f of LADRC is 25% and σ_f of multi-ESO is 5%, and t_{sf} of LADRC is 2.5 s and that of multi-ESO is 1.8 s. In the case of a large interference first, the standard LADRC control is slightly better. Here, the e_f of LADRC and multi-ESO control is 0, the σ_f of LADRC is 50% and σ_f of multi-ESO is 55%, and t_{sf} of LADRC is 1.5 s and that of multi-ESO is 1.2 s. The reverse oscillation after stuck failure was injected may be due to the unexpected coupling calculation of the multi-ESO based control. Nevertheless, the degradation of the control quality of the multi-ESO based control strategy during stuck failure C is acceptable.



(a) First 10 degrees and then 20 degrees



(b) First 20 degrees and then 10 degrees.

Fig. 11. Simulation results in failure state C

(4) Reliability Analysis

Based on the responses of the LADRC and multi-ESO control, the motion reliability of both methods can be calculated by equations (15)-(19), and the calculation results are shown in TABLE 2.

In the table, as in all the situations, the system could handle a disturbance of 20 degrees, and the steady-state errors are zero. The sta-

Table 2. Reliability Calculation of LADRC and Multi-ESO

Situation	Method	Overshoot, σ (%)	Response Time, ts (s)	Transient Reliability, R_t	Time Coefficient, t_R
Normal (20°)	LADRC	12.5	1.5	/	/
	m-ESO	15	1.5	/	/
Failure state A (20°)	LADRC	32.5	2.5	0.77	0.6
	m-ESO	12.5	1.8	1	0.83
Failure state B (20°)	LADRC	60	2.6	0.46	0.58
	m-ESO	15	1.6	1	0.94
Failure state C (10° first)	LADRC	25	2.5	0.86	0.6
	m-ESO	5	1.8	1	0.83
Failure state C (20° first)	LADRC	50	1.5	0.57	1
	m-ESO	55	1.2	0.53	1
Method	**Stability Reliability R_s**	**Steady State Reliability R_r**	**Transient Reliability R_t**	**Time Coefficient t_R**	**Motion Reliability R_m**
LADRC	1	1	0.46	0.58	0.2668
m-ESO	1	1	0.53	0.83	0.4399

bility reliability, R_s, and the steady state reliability, R_r, are 1 for both control methods. The transient reliability, R_t, and time coefficient, t_R, of both control methods take the minimum values under these fault conditions. From the calculation results in TABLE 2, it is obvious that the reliability of the proposed multi-ESO control is better than that of the LADRC under the fault conditions.

According to the system simulation results under normal conditions and several failure conditions, the flexible control output of the multi-ESO based control strategy gives it excellent performance when the standard LADRC performance is poor. Although in the cases of normal conditions or some simple failures, the control performance of the multi-ESO based control method is slightly worse than that of the standard LADRC. The proposed multi-ESO based control strategy will enable the quadruped robot level adjustment control system to better cope with the leg stuck failures and maintain its anti-interference ability.

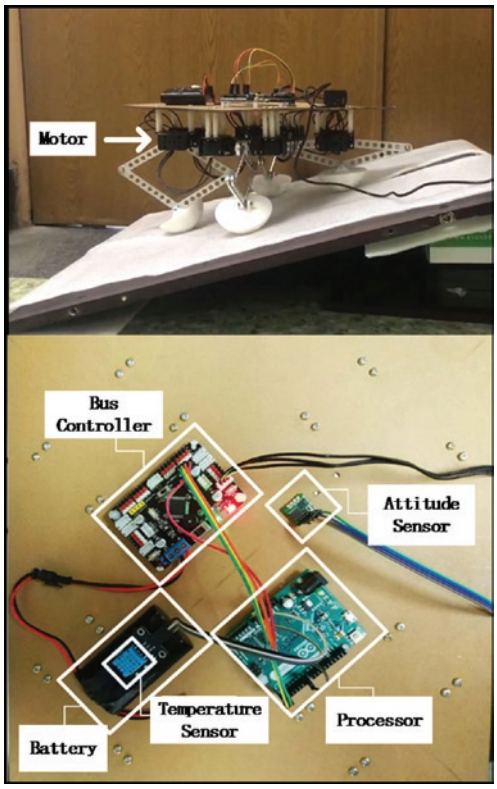


Fig. 12. Experimental setup

5.2. Pro-type Design and Experimental Validation

To show the effectiveness of the proposed multi-ESO based LADRC control solution in real time with leg failures, an experimental validation with the same structure introduced in Section 4 was set up in the laboratory as shown in Fig. 12. The tilt angle of the robot platform is detected by a JY-61 attitude sensor. The processor adopted was an Arduino Leonardo circuit board with an Atmega32u4 microcontroller. Eight ZX3615 motors were driven by a bus controller

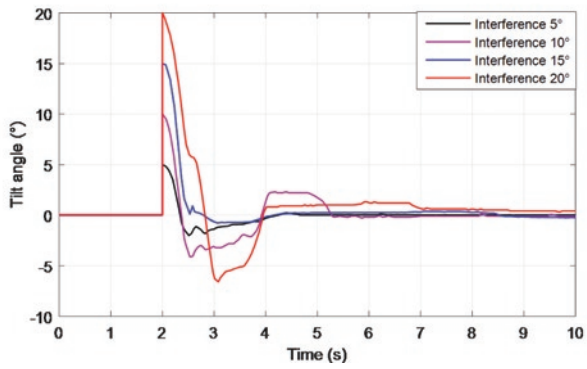


Fig. 13. Experimental results of the multi-ESO based control strategy under normal conditions

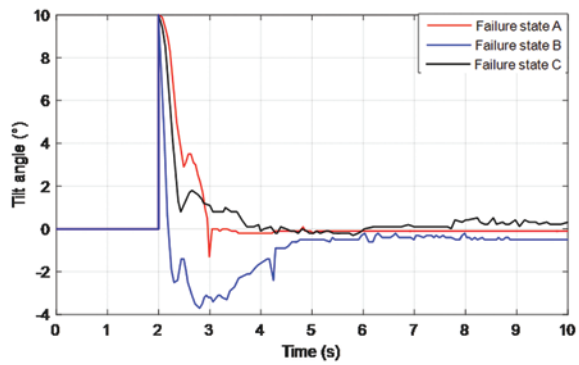


Fig. 14. Experimental results of the multi-ESO based control strategy under failure conditions

powered by batteries. The output voltage of the batteries is 12 V. To monitor the battery status, a temperature sensor was also added.

A robot using the multi-ESO based control algorithm was placed on slopes with inclinations of 5, 10, 15, and 20 degrees. The experimental curves of the robot platform tilt angles without leg failures were obtained by the computer program and are shown in Fig. 13. It can be seen that the system using the multi-ESO based control strategy can effectively overcome an external disturbance of no more than 20 degrees. The system altitude can quickly return to the target value with little overshoot, and the response time is short. Under large angle interference, the system response time is slightly longer, and there is a small static error, which is mainly due to hardware limitations. The results of the physical experiment are basically consistent with the simulation results.

The quadruped robot was placed on a 10 degree slope, and the leg stuck failures were injected by modifying the servo control command to obtain the system tilt angle output curves under failure conditions. Experimental results of the proposed multi-ESO based control strategy in failure states A, B and C are shown in Fig. 14. For failure state C, only the response to the second interference is shown. In the cases of failure state B and failure state C, there are slight jitters and static error in the system. According to the results, the robot could handle the disturbance with a leg stuck at different positions, which validates the effectiveness of the proposed control strategy.

6. Conclusion

In this paper, a multi-ESO based LADRC control strategy for the level adjustment control system of a quadruped robot is proposed and analyzed. By treating multiple legs as parallel control objects, the parameters of the ESO designed for each leg can be the same, which will not increase the number of parameters used in the LADRC control method. The control signals received by each leg can be dynamically adjusted due to the different state observations and estimates of each leg. For the reliability analysis, we took the condition of leg fault occurring into account and proposed a multidimensional reliability evaluation method for control systems. Control performance and reliability of the proposed multi-ESO based control strategy and the standard LADRC method were systematically simulated and compared. An experimental quadruped robot platform was established in the laboratory to show the effectiveness of the proposed multi-ESO based control solution in actual situations. The results illustrate that the proposed multi-ESO-based control method has the advantages of easy parameter tuning, good robustness, and strong ability to cope with disturbance and failure conditions. The proposed multi-ESO control method has more flexible control output to actuators, which can improve the reliability of the control system.

The main contributions of this paper are as follows:

(1) A multi-ESO based LADRC strategy for a level adjustment control system of a quadruped robot is proposed that could improve the fault tolerance of the LADRC under faults.

(2) A multi-dimensional reliability evaluation method for control systems was provided and used to analyze and evaluate the reliability of the proposed multi-ESO method under fault conditions. The steady state, transient and time characteristics of the control system can be comprehensively considered by this method, and the reliability of the control system response can be evaluated.

For future work, we would try to apply the multi-ESO control method to more objects. The synthesized part can be replaced by fuzzy modules or other intelligent methods.

References

1. Castaneda L, Luviano-Juarez A, Chairez I. Robust Trajectory Tracking of a Delta Robot Through Adaptive Active Disturbance Rejection Control. IEEE Transactions on Control Systems Technology 2015; 23(4): 1387-1398, https://doi.org/10.1109/TCST.2014.2367313.
2. Christensen D, Schultz U, Stoy K. A distributed and morphology-independent strategy for adaptive locomotion in self-reconfigurable modular robots. Robotics & Autonomous Systems 2013; 61(9): 1021-1035, https://doi.org/10.1016/j.robot.2013.05.009.
3. Cong Z. Distributed ESO based cooperative tracking control for high-order nonlinear multi-agent systems with lumped disturbance and application in multi flight simulators systems. Isa Transactions 2018; 74: 217-228, https://doi.org/10.1016/j.isatra.2018.01.020.
4. Cui J, Zeng S, Ren Y, Chen X, Gao Z. On the robustness and reliability in the pose deformation system of mobile robots. IEEE Access 2018, https://doi.org/10.1109/ACCESS.2018.2835836.
5. Cully A, Clune J, Tarapore D, Mouret J. Robots that can adapt like animals. Nature 2015; 521(7553): 503-507, https://doi.org/10.1038/nature14422.
6. Du B, Wu S, Han S, Cui S. Application of Linear Active Disturbance Rejection Controller for Sensorless Control of Internal Permanent-Magnet Synchronous Motor. IEEE Transactions on Industrial Electronics 2016; 63(5): 3019-3027, https://doi.org/10.1109/TIE.2016.2518123.
7. Falconi G, Heise C, Holzapfel F. Fault-tolerant position tracking of a hexacopter using an Extended State Observer. International Conference on Automation, Robotics and Applications. IEEE 2015; 550-556, https://doi.org/10.1109/ICARA.2015.7081207.
8. Gonzalez-Prieto I, Duran M, Barrero F. Fault-Tolerant Control of Six-Phase Induction Motor Drives With Variable Current Injection. IEEE Transactions on Power Electronics 2017; 32(10): 7894-7903, https://doi.org/10.1109/TPEL.2016.2639070.
9. Han J. From PID to active disturbance rejection control. IEEE Transactions on Industrial Electronics 2009; 56(3): 900-906, https://doi.org/10.1109/TIE.2008.2011621.
10. Herbst G. Practical Active Disturbance Rejection Control: Bumpless Transfer, Rate Limitation, and Incremental Algorithm. IEEE Transactions on Industrial Electronics 2016; 63(3): 1754-1762, https://doi.org/10.1109/TIE.2015.2499168.
11. Huang Y, Xue W. Active disturbance rejection control: Methodology and theoretical analysis. ISA Transactions 2014; 53(4): 963-976, https://doi.org/10.1016/j.isatra.2014.03.003.
12. Kommuri S, Defoort M, Karimi H, Veluvolu K. A Robust Observer-Based Sensor Fault-Tolerant Control for PMSM in Electric Vehicles. IEEE Transactions on Industrial Electronics 2016; 63(12): 7671-7681, https://doi.org/10.1109/TIE.2016.2590993.
13. Li D, Ding P, Gao Z. Fractional active disturbance rejection control. ISA Transactions 2016; 62: 109-119, https://doi.org/10.1016/j.isatra.2016.01.022.
14. Li D, Li C, Gao Z, Jin Q. On active disturbance rejection in temperature regulation of the proton exchange membrane fuel cells. Journal of Power Sources 2015; 283: 452-463, https://doi.org/10.1016/j.jpowsour.2015.02.106.
15. Li J, Xia Y, Qi X, Gao Z. On the Necessity, Scheme and Basis of the Linear-Nonlinear Switching in Active Disturbance Rejection Control. IEEE Transactions on Industrial Electronics 2017; 64(2): 1425-1435, https://doi.org/10.1109/TIE.2016.2611573.
16. Liu F, Li Y, Cao Y, She J, Wu M. A Two-Layer Active Disturbance Rejection Controller Design for Load Frequency Control of Interconnected

Power System. IEEE Transactions on Power Systems 2016; 31(4): 3320-3321, https://doi.org/10.1109/TPWRS.2015.2480005.
17. Ma X, Sun F, Li H, He B. Neural-network-based integral sliding-mode tracking control of second-order multi-agent systems with unmatched disturbances and completely unknown dynamics. International Journal of Control, Automation and Systems 2017; 15(4): 1925-1935, https://doi.org/10.1007/s12555-016-0057-z.
18. Madonski R, Kordasz M, Sauer P. Application of a disturbance-rejection controller for robotic-enhanced limb rehabilitation trainings. ISA Transactions 2014; 53(4): 899-908, https://doi.org/10.1016/j.isatra.2013.09.022.
19. Marina S, Franc N, Gregor P. Sensors in proactive maintenance - a case of LTCC pressure sensors. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(2): 267-272, https://doi.org/10.17531/ein.2018.2.12.
20. Ramos G, Cortes-Romero J, Coral-Enriquez H. Spatial observer-based repetitive controller: An active disturbance rejection approach. Control Engineering Practice 2015; 42: 1-11, https://doi.org/10.1016/j.conengprac.2015.05.002.
21. Ran M, Wang Q, Dong C. Stabilization of a class of nonlinear systems with actuator saturation via active disturbance rejection control. Automatica 2016; 63: 302-310, https://doi.org/10.1016/j.automatica.2015.10.010.
22. Rymarczyk T, Klosowski G. Innovative methods of neural reconstruction for tomographic images in maintenance of tank industrial reactors. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(2): 261-267, https://doi.org/10.17531/ein.2019.2.10.
23. Song Y, Guo J. Neuro-Adaptive Fault-Tolerant Tracking Control of Lagrange Systems Pursuing Targets With Unknown Trajectory. IEEE Transactions on Industrial Electronics 2017; 64(5): 3913-3920, https://doi.org/10.1109/TIE.2016.2644606.
24. Su X, Liu X, Song Y. Fault-Tolerant Control of Multi-area Power Systems via a Sliding-Mode Observer Technique. IEEE/ASME Transactions on Mechatronics 2018; 23(1): 38-47, https://doi.org/10.1109/TMECH.2017.2718109.
25. Shen Q, Wang D, Zhu S, Poh E. Robust Control Allocation for Spacecraft Attitude Tracking Under Actuator Faults. IEEE Transactions on Control Systems Technology 2017; 25(3): 1068-1075, https://doi.org/10.1109/TCST.2016.2574763.
26. Wu D, Chen K. Limit cycle analysis of active disturbance rejection control system with two nonlinearities. ISA Transactions 2014; 53(4): 947-954, https://doi.org/10.1016/j.isatra.2014.03.001.
27. Xiao B, Yin S. Velocity-Free Fault-Tolerant and Uncertainty Attenuation Control for a Class of Nonlinear Systems. IEEE Transactions on Industrial Electronics 2016; 63(7): 4400-4411, https://doi.org/10.1109/TIE.2016.2532284.
28. Xu C, Li J, Zhang P, Mu L, Si X. ESO-based fault diagnosis and fault-tolerant for incipient actuator faults. Control and Decision Conference. IEEE 2013; 4359-4363.
29. Xue W, Bai W, Yang S, Song K, Huang Y, Xie H. ADRC With Adaptive Extended State Observer and its Application to Air-Fuel Ratio Control in Gasoline Engines. IEEE Transactions on Industrial Electronics 2015; 62(9): 5847-5857, https://doi.org/10.1109/TIE.2015.2435004.
30. Xue W, Huang Y. On performance analysis of ADRC for a class of MIMO lower-triangular nonlinear uncertain systems. ISA Transactions 2014; 53(4): 955-962, https://doi.org/10.1016/j.isatra.2014.02.002.
31. Yang X, Cui J, Lao D, Li D, Chen J. Input Shaping enhanced Active Disturbance Rejection Control for a twin rotor multi-input multi-output system (TRMS). ISA Transactions 2016; 62: 287-298, https://doi.org/10.1016/j.isatra.2016.02.001.
32. Yang Y, Huang H, Liu Y, Zhu S, Peng W. Reliability analysis of electrohydraulic servo valve suffering common cause failures. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(3): 354-359.
33. Yu B, Paassen A. Simulink and bond graph modeling of an air-conditioned room. Simulation Modelling Practice & Theory 2004; 12(1): 61-76, https://doi.org/10.1016/j.simpat.2003.12.001.
34. Zhou Y, Huang Z, Liu W, Li H, Liao H. A distributed ESO based cooperative current-sharing strategy for parallel charging systems under disturbances. IEEE Energy Conversion Congress & Exposition 2016, https://doi.org/10.1109/ECCE.2016.7854675.

Jingjing CUI
Yi REN
Binghui XU
Dezhen YANG
Shengkui ZENG
School of Reliability and Systems Engineering
Beihang University
No.37 Xueyuan RD. Haidian, 100191, Beijing, China

E-mails: cuijingjing@buaa.edu.cn, renyi@buaa.edu.cn, xubinghui@buaa.edu.cn, dezhenyang@buaa.edu.cn, zengshengkui@buaa.edu.cn

Article citation info:
SOBASZEK Ł, GOLA A, ŚWIĆ A. Time-based machine failure prediction in multi-machine manufacturing systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 52–62 http://dx.doi.org/10.17531/ein.2020.1.7.

Łukasz SOBASZEK
Arkadiusz GOLA
Antoni ŚWIĆ

TIME-BASED MACHINE FAILURE PREDICTION IN MULTI-MACHINE MANUFACTURING SYSTEMS

ALGORYTM WSPARCIA STRATEGII TBM W WIELOMASZYNOWYCH SYSTEMACH WYTWÓRCZYCH*

The execution of production processes in real manufacturing systems is associated with the occurrence of numerous disruptions, which predominantly revolve around technological machine failure. Therefore, various maintenance strategies are being developed, many of which tend to emphasise effective preventive measures, such as the Time-Based Maintenance (TBM) discussed in this paper. Specifically, this publication presents the time-based machine failure prediction algorithm for the multi-machine manufacturing environment. The Introduction section outlines the body of knowledge related to typical strategies applied in maintenance. The next part describes an approach to failure prediction that treats processing times as makespan and is followed by highlighting the key role of historical data in machine failure management, in the subsequent section. Finally, the proposed time-based machine failure prediction algorithm is presented and tested by means of a two-step verification, which confirms its effectiveness and further practical implementation.

Keywords: *production system, maintenance, reliability, machine failure, prediction, Time-Based Maintenance.*

Realizacja procesów produkcyjnych w rzeczywistych systemach wytwórczych wiąże się z występowaniem wielu zakłóceń, do których zalicza się głównie awarie maszyn technologicznych. W związku z tym obserwowany jest rozwój różnorodnych strategii utrzymania ruchu. Coraz większy nacisk kładziony jest na efektywne działania prewencyjne, do których zalicza się także działania określone w czasie (ang. Time-Based Maintenance – TBM). W niniejszej publikacji zaprezentowano algorytm predykcji awarii maszyn w wielomaszynowych systemach wytwórczych wspierający prewencyjne utrzymanie ruchu. Na wstępie omówiono zagadnienia związane z typowymi strategiami stosowanymi w obszarze UR. Ponadto omówiono tematykę predykcji awarii, zwracając uwagę na ujęcie czasu pracy maszyny jako czasu trwania, a także kluczową rolę wykorzystania danych historycznych dotyczących awarii maszyn. Następnie zaprezentowano proponowany algorytm predykcji wspierający działania określone w czasie. Prezentowane prace zakończono dwuetapową weryfikacją proponowanej metody, która potwierdziła jej skuteczność oraz zasadność wykorzystania.

Słowa kluczowe: *system produkcyjny, utrzymanie ruchu, niezawodność, awarie maszyn, predykcja, Time-Based Maintenance.*

1. Introduction

The reality of the production environment is inseparably connected with disruptions, which negatively impact the executed processes, thus leading to disorganisation [14]. The key uncertainty factors include the occurrence of technological machine failure. From the practical point of view, prediction of failure times is an issue of fundamental importance, as it enables implementing preventive activities in a way that does not interfere with the current production process. Failure time prediction is frequently in use in Time-Based Maintenance (TBM), and in response to the growing demand, specialised IT solutions aimed to support this strategy are developed [5, 16, 37]. It is crucial that these tools employ effective prediction algorithms, drawing from reliable historical data and thus providing the basis for a reliable analysis of machine failure and proper adjustment of maintenance activities [6, 13, 40].

The literature analysis shows that numerous studies have been devoted to the prediction of disruption in the production process. Those studies primarily concern the development of effective methods for countering failure, as well as absorb their impact [3, 33]. Preventive activities frequently correspond with the principles of Time-Based Maintenance [13, 25], as well as activities representing Conditioned-Based Maintenance [1, 30]. The development of scenarios and operational strategies is also a very popular trend [26, 27, 34, 35, 39].

Failure prediction methods proposed in the literature are categorised into several groups:
- methods based on probability distribution,
- methods using typical performance indicators,
- alternative failure prediction methods,
- methods based on real data.

The vast majority of the solutions proposed in the literature are based on probability distribution analysis [8, 15, 24, 2], which considers typical distributions and their combinations, such as: uniform distribution [17, 2], normal distribution [8] or exponential distribution [24, 30]. The primary purpose of distribution analysis is to define the

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

time of failure occurrence. Solutions based on combinations of typical distributions are also proposed in the literature, for instance, in the 2010 study [15] the authors propose combinations of normal, triangular and exponential distributions to describe the problem of failure occurrence. Admittedly, most of the proposed solutions consider the problem in a purely theoretical manner, and as such, disregard the critical aspect of prediction: the use of historical data on machine failure rate. Furthermore, researchers fail to provide a sufficient justification for a given probability distribution selection.

Another trend visible in the literature is employing key performance indicators (KPIs) used in maintenance for failure prediction, such as:

– *Mean Time To Failure (MTTF)*,
– *Mean Time Between Failures (MTBF)*,
– *Mean Time To Repair (MTTR)*.

The KPIs listed above are employed in numerous studies [9, 12, 21, 20], predominantly directly, in other cases indirectly – as estimators for the purpose of Weibull distribution [21]. In research, the authors follow predefined scenarios and the indicators are specified from preset ranges, which ensures that the failure events occur at a desired frequency (frequently or rarely) and are eventually analysed from the perspective of the consequences of failure occurrence [12]. Sometimes the use of KPIs is supported by the use of appropriate statistical methods [30]. The use of methods applying performance indicators typical for maintenance is substantiated by the fact that these parameters provide large amounts of information on the technological machines in use. Nevertheless, the acquisition and use of parameters in question is largely in the theoretical domain: the published studies fail to perform verification of the proposed solutions with the real data on machine failure rates [9, 20].

With respect to alternative methods of failure prediction, several solutions are particularly worth highlighting, *e.g.* the methods in which all machine failures are accumulated into one and evaluated by means of the MTTR and MBL (Machine Breakdown Level) parameters [18], the methods where the failure rate is determined from the analysis of the machine loading time distributions [31], those in which the prediction of machine failure is carried out with the application of artificial neural networks [4], or the well-established time series models [38]. During the verification of the proposed solutions, however, test data is employed, which, furthermore, stems from the use of simplifying assumptions adopted by the researchers.

In the works of Davenport *et al.* and Kempa *et al.* [8, 19], the authors note that performing computations on actual sets of process data is of paramount importance. These suggestions represent a novelty approach to failure prediction. They point out the necessity to develop methods focusing on the practical use of historical data on technological machine failure. Although studies implementing such solutions may be found in the specialist literature, their number is still negligible [33]. Nonetheless, they represent a clear trend in the area of failure prediction.

Despite the fact that several methods have been proposed, no solutions towards the practical use of historical data on the failure of technological machines have yet been developed. In addition, in the production environment the typical *modus operandi* is to propose implementation of extensive and high-priced monitoring systems, while in the field of TBM strategies, the data is obtained from all maintenance departments. Therefore, this study provides a novel approach to machine failure prediction in multi-machine manufacturing systems that employs an algorithm performing an in-depth, elaborate analysis of actual production data, thus enabling the prediction of future machine breakdowns and implementation of effective preventive measures. This method constitutes an alternative to those characterised in the preceding paragraphs as it makes use of data obtained from maintenance services to achieve the intended objective – identification of the potential moment of failure. The innovation of our method consists in its incorporation of elements of survival analysis theory in technological machine failure analysis enabling statistical inference based on historical data.

2. Failure prediction with elements of processing times analysis

2.1. Machining times as duration

In its essence, failure prediction is the determination of the time and degree of certainty for the occurrence of failure of a given technological machine; to this end, elements of Survival Analysis, also referred to as Duration Analysis [11, 23], may be put to use.

When employing Duration Analysis it is essential to precisely specify the essence of the studied process, which should meet the following conditions [11]:

1. Changes to the analysed unit are made between discrete states.
2. Changes of states occur at any time and are not fixed in time.
3. Changes are reversible or irreversible (relative to the form of the process).
4. Changes are predetermined by the current state of the process.
5. Certain factors affect the process – the analysis enables their detection.

Considering these determinants of the Survival Analysis, it appears that technological machine failure is a process that meets these requirements. Machine failure can occur at any time and is a change between two states – the functioning and breakdown. In addition, damage to the machine is a reversible change – once repaired, it returns to its original state, being defined by the state in which the device is. There are also a number of factors that can affect the process under scrutiny and can be identified by means of Duration Analysis [36]. In the case of machine there is a need to consider the duration time as a time of undisturbed machine operation. In the consequence, the failure time of machine can be determined. An additional advantage of this technique is the ability to determine failure patterns (time characteristics of failures), especially when the historical data do not allow the use of typical inference techniques [33].

Let T be a non-negative random variable representing the time of failure (duration) of the technological machine, whose value is in the range $(0; \infty)$. In addition, $f(t)$ is a function of probability density, where $t > 0$ and $F(t)$ is a cumulative distribution function of the random variable, T – a non-decreasing function that indicates that the object will experience the event in time $(0; t]$:

$$F(t) = P(T < t) \,. \tag{1}$$

Based on the cumulative distribution function $F(t)$, the survival function $S(t)$ can be defined as:

$$S(t) = 1 - F(t) = P(T \geq t) = \int_t^{\infty} f(s)ds \,, \tag{2}$$

which gives the probability of undisturbed machine work until t. It, furthermore, determines the probability that a failure will not occur until t. The selected function is an ideal solution for the determination of patterns of correct machine operation and, as a consequence, also its failure. The survival and cumulative function are shown in Figure 1.

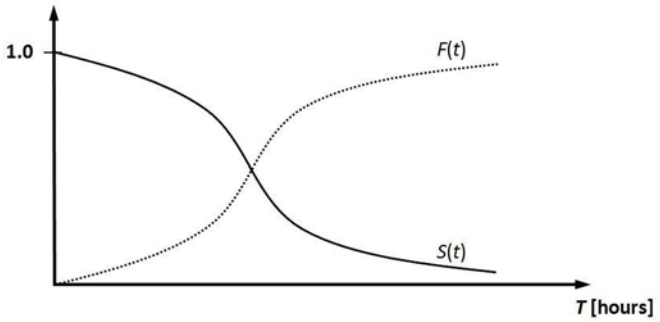


Fig. 1. Cumulative distribution function F(t) and survival function S(t)

In order to determine the particular functions presented above, appropriate historical data describing the failure of the technological machine should be obtained and incorporated in the models. Their analysis provides a great amount of critical information that can be used in the further prediction process.

2.2. The use of historical data

To determine the failure characteristics, it is necessary to define the suitable data source, *i.e.* production maintenance departments – since these cells collect the information in question [3, 10]. The data on the history of maintenance and repair of technological machines in manufacturing enterprises are most commonly recorded by means of the following solutions:

- paper documentation – typically in the form of Maintenance Cards and Service Books,
- IT software coupled with dedicated spreadsheets (Fig. 2),
- data acquisition directly from technological machines, using SCADA (Supervisory Control And Data Acquisition) and MES (Manufacturing Execution Systems).

Fig. 2. An example of service data recorded in a computer spreadsheet

All of the data collection methods above share a common feature – each provides information that, when properly processed, can be employed in Survival Analysis for the prediction of machine failure.

The data contained in the documentation are historical failure times. For a given technological machine M_j, they are given as T_{Mj}:

$$T_{Mj} = \{t_1, t_2, ..., t_n\} \text{ [hours]}, \tag{3}$$

where: t_i – i-th time of failure.

An example dataset for M_1 historical failure times is expressed by:

$$T_{M1} = \{4, 8, 20, 16, 10, 28, 43, 15, 24, 2, ...\} \text{ [hours]}.$$

The use of data contained in relevant datasets T_{Mj} enables the determination of potential failure times of a given machine, saved in dataset FT_{Mij}:

$$FT_{Mj} = \{ft_{Mj1}, ft_{Mj2}, ..., ft_{Mjn}\}, \tag{4}$$

where: ft_{Mji} – failure time of machine j,
j – the number of the considered machine.

For each time ft_{Mji} the probability of failure is given in the set P_{Mj}.

$$P_{Mj} = \{p_{Mj1}, p_{Mj2}, ..., p_{Mjn}\}, \tag{5}$$

where: p_{Mji} – the probability of machine failure j, given that:

$$\bigwedge_{ft_{Mij} \neq 0} p_{Mij} \neq 0 .$$

Therefore, the result of the prediction will be the pairs (p_{Mji}, ft_{Mji}) that define the probability and the failure time of machine M_j.

2.3. The proposed time-based machine failure prediction algorithm

In order to predict the probability of failure and the time of failure, a four-step algorithm was developed to analyse and properly implement the collected repair history data.

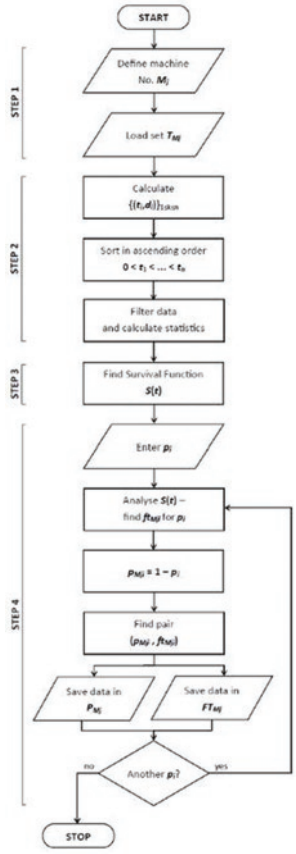


Fig. 3. Failure prediction algorithm

Step 1 of the proposed algorithm defines the machine for which the prediction process is carried out, as well as acquires the historical data from in the set T_{Mj} (Fig. 3).

At step 2, the imported data are saved: the failure times of machine M_j by means of an appropriate sequence:

$$\{(t_i, d_i)\}_{1 \le k \le n}, t_i \in T_{Mj} \tag{6}$$

where: t_i – the time between successive failures,
d_i – number of cases.

In addition, at this step the data is arranged in an increasing order $\{t_i\}_{1 \le k \le n}$:

$$0 < t_1 < t_2 < ... < t_n, \tag{7}$$

Subsequently, the acquired data are filtered and outliers (representing atypical values) removed (Fig. 4). Then, the basic statistics for the collected data (minimum, maximum, average deviation, quartile range) are determined.

Step 3 is crucial for the inference process because it is at this stage that the survival function, characterising the considered failure process of the analysed machine, is determined. By ordering machine failures according to the increasing occurrence times and by determining the number of cases for each such occurrence, the survival function of a given process is determined. The obtained function conveniently determines duration patterns (failure occurrence) and allows to determine failure characteristics of the defined machine. The application of Kaplan-Meier estimation, on the other hand, produces the survival function, determined from the relationship:

$$\hat{S}(t) = \begin{cases} 1, & for\ t < t_1 \\ \prod_{t_i \le t} \frac{r_i - d_i}{r_i}, & for\ t_1 < t \end{cases} \tag{8}$$

where: r_i – the number of all breakdowns, given by:

$$r_i = \sum_{j=i}^{k} d_j\ . \tag{9}$$

Subsequently, the survival function is determined, which allows to determine (with defined probability level) the undisturbed machine operation times (Fig. 5).

The determined survival function is implemented at step 4, where the obtained results serve to determine the elements of searched sets:

- potential times of machine failure FT_{Mj},
- probability of machine failure P_{Mj}.

Fig. 6 shows the principles of statistical inference based on the survival function. Predictions of failure times ft_{Mji} are determined for specified probability levels p_i.

Since the probability of undisturbed machine operation (p_i) is determined from the survival function, therefore, machine failure probability p_{Mji} is given by:

$$p_{Mji} = 1 - p_i, \tag{10}$$

where: p_{Mji} – machine failure probability,
p_i – undisturbed machine operation probability.

Determining the searched machine failure probability p_{Mji} enables the determination of the searched ft_{Mji}, and, consequently, determining the pairs (p_{Mji}, ft_{Mji}). The calculated data are collected in sets P_{Mji} and

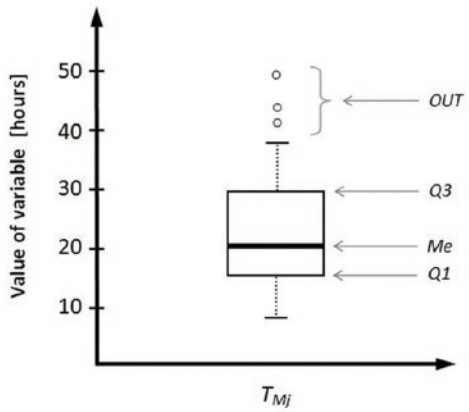


Fig. 4. Box chart for sample data (Me – median, Q1 and Q3 – quartiles 1 and 3, OUT – outliers)

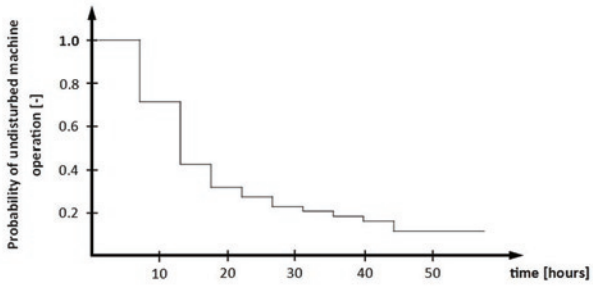


Fig. 5. An example of Survival Function determined using Kaplan-Meier estimation

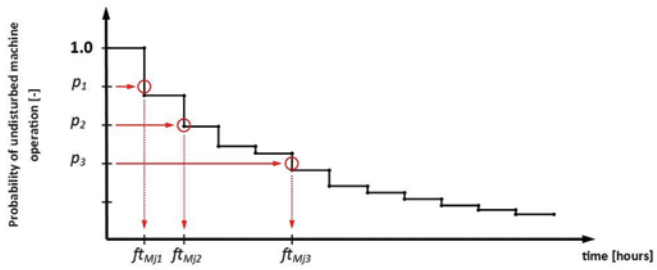


Fig. 6. Determining the failure time based on the adopted value of survival probability

FT_{Mji}. Step 4 is iterative and is, therefore, repeated depending on the user's decision regarding the number of probability levels to consider. The implementation of the algorithm should be repeated for other technological machines whose failure rate is investigated.

3. Experimental verification of the proposed algorithm

3.1. Data used in verification

The step preceding the model verification, presented below, was the acquisition and implementation of data describing the characteristics of the executed technological processes and the failure rate of technological machines. As mentioned before, the investigations reported in this study were based on actual production data, which specifically consisted of 12 production tasks performed at 12 work stations, arranged in manufacturing cells. The prevailing manufacturing process carried out in production is subtractive machining. Table 1 below lists technological processes at selected production jobs.

Table 1. Examples of technological processes contained in production data

Product No. (job)	Operation No.	Workstation	Operation	ts_{ij} [hours]	to_{ij} [hours]
1	10	Laser1	Cutting sheets	0.25	0.042
	20	Laser2	Laser-cutting pipes and profiles	0.20	0.017
	30	CNC Press	Edge bending	0.13	0.018
	40	Drilling machine	Drilling holes	0.17	0.017
	50	Metalworking	Metalworking	0.08	0.017
	60	MIG welder	MIG welding	0.13	0.092
3	10	Laser2	Laser-cutting pipes and profiles	0.15	0.005
	20	CNC band saw	Band-saw cutting	0.10	0.008
	30	Milling machine	Milling	0.27	0.050
	40	Drilling machine	Drilling holes	0.17	0.017
	50	Metalworking	Metalworking	0.08	0.033
	60	MIG welder	MIG welding	0.13	0.033
	70	Turning lathe	Turning	0.33	0.092
5	10	Laser1	Laser-cutting metal sheets	0.27	0.012
	20	Metal shearing machine	Metal shearing	0.10	0.004
	30	CNC band saw	Band-saw cutting	0.10	0.017
	40	CNC press	Edge bending	0.17	0.025
	50	Drilling machine	Tapping	0.13	0.100
	60	Metalworking	Metalworking	0.08	0.033
	70	MIG welder	MIG welding	0.13	0.033
	80	Turning lathe	Turning	0.33	0.108

The actual data used in the verification process were employed in the following scope:

- technological machine failure data were used as input data for the prediction algorithm verification,
- data on executed manufacturing processes were used in simulation tests to assess the effectiveness and validity of the proposed algorithm considering real production conditions (including technological machinery failure).

3.2. Failure time prediction

The proposed algorithm was verified by means of an appropriate script compiled in a programming language R. The successful verification was followed by the use of the historical data in the process of statistical inferring with respect to the potential breakdown times of machines at particular workstations. The machines constituting the stock of the machine tools were labelled as follows:

- Laser 1 – machine M_1,
- Laser 2 – machine M_2,
- CNC press – machine M_3,
- CNC band saw – machine M_4,
- Metalworking station – machine M_5,
- MIG welder – machine M_6,
- TIG welder – machine M_7,
- Drilling machine – machine M_8,
- Milling machine – machine M_9,
- Turning lathe – machine M_{10},
- Metal shearing machine – machine M_{11},
- Punching machine – machine M_{12}.

In the paragraphs below, the exemplary execution of the verification process is presented for machine M_6, in which case the historical data included 121 observations

Prior to the initiation of the prediction process, the prepared script was fed with appropriate commands – preparing the software working environment; this was followed by specifying the machine number and importing the data from the *.CSV file. By importing the data into the set T_{M6}, (the variable) stored in the workspace, facilitated sorting the considered observations in ascending order, as well as filtering the data by means of the box plots (Fig. 7). In addition, basic statistics were determined (Fig. 8).

The key step of our failure prediction algorithm is the determination of the survival function, $\hat{S}(t)$, with the application of Kaplan-Meier estimation, which was enabled by including the "SURVIVAL" library in the script. A further course of the step function was

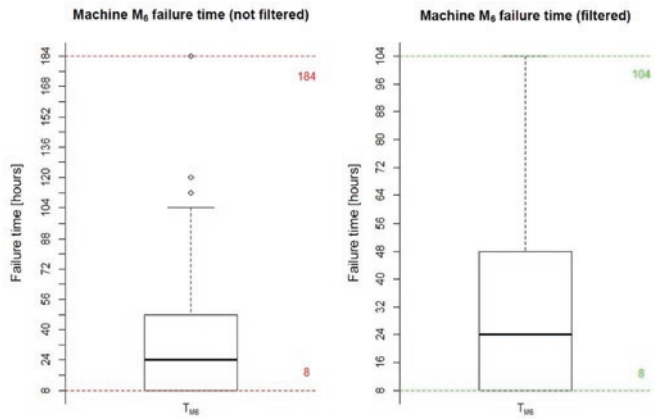


Fig. 7. Box plots – before and after data filtering

calculated automatically from the produced observation sequences. The result was a survival function in the form of a stepped curve at 95% confidence.

Determining the course of the searched function, $\hat{S}(t)$, triggers the next step of the algorithm: the prediction of the failure time of the considered machine at the defined probability level (Fig. 9). As the probability of undisrupted machine operation can also be read from the chart, an additional legend with explanations was generated. In the

```
"Machine M6 failure rate - set TM6 - basic statistics:"
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
8.00    8.00   24.00   32.48   48.00  104.00
```

Fig. 8. Basic statistics generated by the developed script

case of calculations for the given machine M_6 (and other machines), the following probability levels were considered:

$$p_1 = 0.75; \quad p_2 = 0.50; \quad p_3 = 0.25.$$

The values of the considered levels have been chosen so as to determine: low, medium and high level of risk of the machine being affected. Therefore:

$$p_{M61} = 1 - p_1 = 0.25; \quad p_{M62} = 1 - p_2 = 0.50; \quad p_{M63} = 1 - p_3 = 0.75;$$

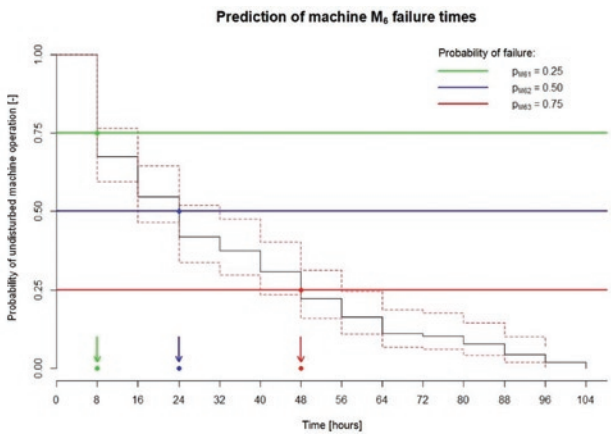


Fig. 9. Failure prediction based on the survival function

In this way, the probability of occurrence and times of potential failures were calculated, and can be expressed as pairs:

$$(p_{M61}, ft_{M61}) = (0.25, 8 \text{ hours}),$$
$$(p_{M62}, ft_{M62}) = (0.50, 24 \text{ hours}),$$
$$(p_{M63}, ft_{M63}) = (0.75, 48 \text{ hours}).$$

As a result, sets $P_{M61} = \{0.25, 0.50, 0.75\}$ and $FT_{M61} = \{8, 24, 48\}$ [hours] were determined.

The proposed algorithm was used to the same extent in other technological machines. Due to the nature of the metalworking workstation (M_5) the prediction process was not carried out. The calculated failure times are given in Table 2.

The results obtained from the executed algorithm were employed in the subsequent part of the verification process, consisting in the simulation of production under technological machinery failure constraint.

3.3. Production simulation under uncertainty

The plan of the study described in this paper assumed the verification of the introduced algorithm in the real production environment in order to validate its applicability under machine failure uncertainty,

Table 2. Technological machine failure times obtained from prediction

	Failure time [hours]		
Machine	$p_{Mj1} = 0.25$	$p_{Mj2} = 0.50$	$p_{Mj3} = 0.75$
M_1	8	16	40
M_2	8	24	32
M_3	8	16	24
M_4	8	24	104
M_5	–	–	–
M_6	8	24	48
M_7	8	16	40
M_8	8	24	48
M_9	8	16	40
M_{10}	8	24	40
M_{11}	8	16	40
M_{12}	8	16	32

which is characteristic of authentic industrial conditions. This was done in a two-stage experiment:

1. Nominal production schedules were produced based on the actual production data. Next, corresponding robust schedules were prepared by implementing service times as indicated by the results of the executed algorithm.
2. The production process was modelled according to the developed schedules and examined to indicate the schedule of the shortest production completion time under the constraint of machine failure.

3.3.1. Scheduling production

Different job scheduling methods to follow at individual workstations were evaluated by means of 4 established dispatching rules:

1. FCFS (First Come First Service).
2. EDD (Earliest Due Date).
3. SPT (Shortest Processing Time).
4. LPT (Longest Processing Time).

It was assumed that the products were made in 50-piece batches, and the objective function of the schedule was to minimise the makespan – $C_{\max}$.

The task scheduling tool employed in the study was LiSA, a software package for solving job scheduling problems typical of real production environments (flow-shop, job-shop or open-shop), which makes use of algorithms in imposing a set of constraints and evaluation criteria [7]. Fig. 10 shows an example schedule solved with the use of LPT dispatching rule.

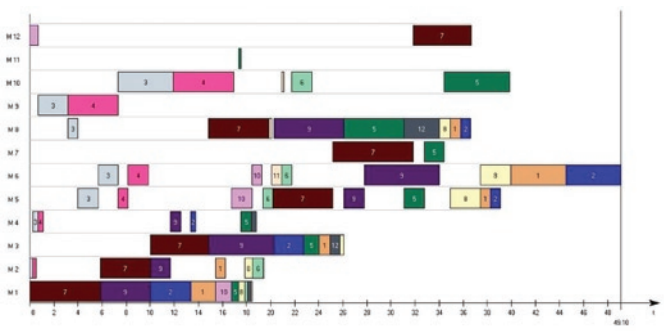


Fig. 10. Nominal schedule – LPT dispatching rule

Potential technological machine failure was accounted for in the schedules by the implementation of service buffers of 0.5 hours, aimed to protect schedules against disruptions and providing the necessary inspection or servicing time. Buffers were incorporated in the schedules in accordance with the indications of the algorithm (Table 2). It was assumed that failure may only occur after the processing time block (processing of jobs). Should there be a technological operation in a given place of the schedule – it would be moved right (immediately after the buffer), thus maintaining the order of tasks indicated in the nominal schedule. An example of a robust schedule with implemented service buffers is shown in Fig. 11 (buffers are represented by white blocks).

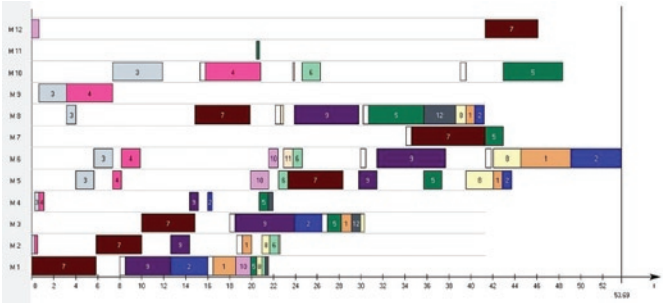


Fig. 11. Production schedule including service buffers

The times of completion of all jobs (makespan) in the nominal and robust schedules are presented in Table 3.

The completion times of all jobs obtained from the test schedules were elongated in every case when time buffers were incorporated. This resulted in the elongation of the objective function C_{max} in each reported case. The average time difference between the nominal and robust schedule amounted to 6.75 h. It may be, therefore, concluded that accounting for technological machine failure causes that the production will extend over approximately one additional shift. Expressed in percentage, the elongation ranged from 8.5% for the robust schedule with the LPT priority rule, to 16.7% for the FCFS schedules. The makespans of particular robust schedules are given in Fig. 12 below.

To evaluate whether the implemented buffers should be incorporated in the schedules, thus leading to the production schedule elongation, the second stage of the verification process was carried out: simulation of production under uncertainty. This step indicated which of the schedules – nominal or robust (produced by the proposed algorithm) – fulfils the objective function, *i.e.* minimisation of completion of all production tasks.

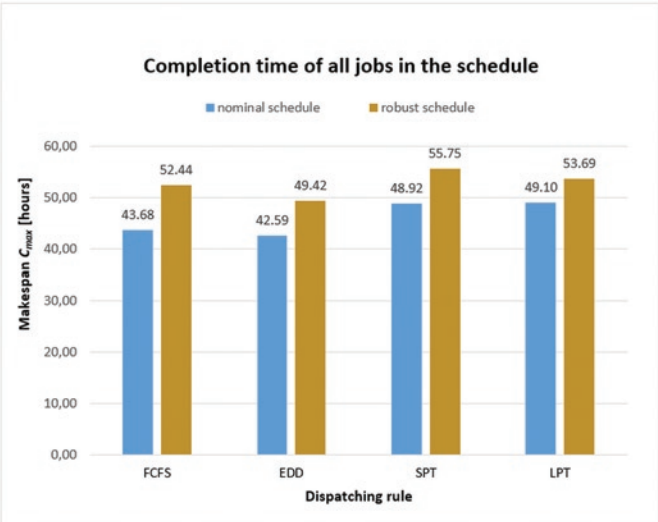


Fig. 12. Makespan C_{max} – completion time of all jobs

Table 3. Obtained values of C_{max}

Dispatching rule	Completion time of all jobs – makespan C_{max} [hours]		
	nominal schedule	robust schedule	elongation [%]
FCFS	43.68	52.44	16.7%
EDD	42.59	49.42	13.8%
SPT	48.92	55.75	12.3%
LPT	49.10	53.69	8.5%

3.3.2. Production simulation under machine failure constraint

The second stage of the experiment was carried out in the Enterprise Dynamics simulation environment, which is one of the leading solutions in simulating various processes. This platform enables representing a range of processes, including production, storage, supply chain management, transport systems, and its capacity for modelling, simulation and visualisation earmarks it for controlling dynamic processes [14, 16, 22]. Putting to use the available elements of the environment, a model was made for the production execution analysis in the considered production system (Fig. 13).

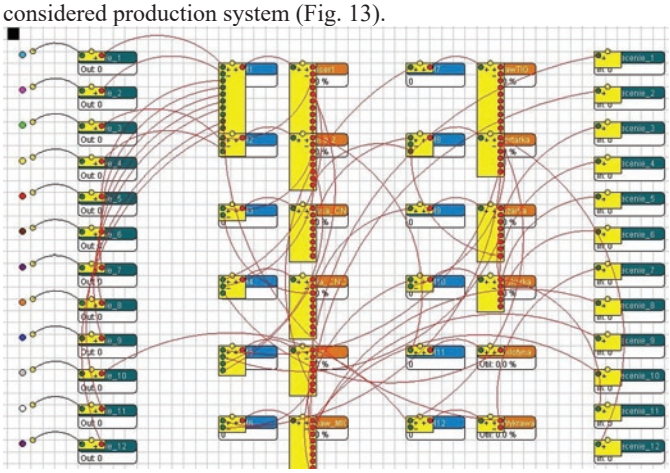


Fig. 13. The production system model developed in the ED environment

Given the failure rate of technological machines, MTTF and MTTR values were defined for each of them, by modifying the properties of a given block. The MTTF parameter values were defined using uniform probability distribution so that the failures occurred at any time – from the commencement of processing jobs on a machine until its completion. The MTTR parameter was determined by gamma distribution, as it was indicated to be the best fitting by the results from the statistical analysis of historical data on machine repair times. The MTTF and MTTR parameters for individual machines are presented in Table 4. Note that due to the ED simulation environment – the times describing the distribution parameters were given in seconds.

The model developed for the purpose of this study included the modification of job orders on particular machines (in accordance with the schedules implementing the particular dispatching rules FCFS, EDD, SPT and LPT).

When assessing the results of simulations, the following stability indicators were used:

– elongation of completion time of all jobs $\Delta C_{\max}$ given by:

$$\Delta C_{\max} = C_{\max} - C'_{\max}, \tag{11}$$

where: $\Delta C_{\max}$ – elongation of completion time of all jobs,
$C_{\max}$ – nominal schedule makespan,

Table 4. Technological machine failure times obtained from the prediction results

	Failure metrics	
Machine	**MTTF**	**MTTR**
M_1	Uniform(0; 66323)	Gamma(3075; 1.62)
M_2	Uniform(0; 31691)	Gamma(2700; 2.07)
M_3	Uniform(0; 57877)	Gamma(2491.8; 2.79)
M_4	Uniform(0; 12013)	Gamma(2773.2; 1.88)
M_5	–	–
M_6	Uniform(0; 85475)	Gamma(3421.2; 2.43)
M_7	Uniform(0; 30024)	Gamma(3352.8; 1.96)
M_8	Uniform(0; 80687)	Gamma(2377.2; 2.45)
M_9	Uniform(0; 24012)	Gamma(2884.8; 1.64)
M_{10}	Uniform(0; 60624)	Gamma(2609.4; 1.85)
M_{11}	Uniform(0; 756)	Gamma(3169.8; 2.16)
M_{12}	Uniform(0; 19800)	Gamma(3015; 1.78)

$C'_{\max}$ – actual (executed) schedule makespan.

– relative elongation of makespan $E_{C\max}$, determined from the relationship:

$$E_{C_{\max}} = \frac{C_{\max}}{C'_{\max}}, \tag{12}$$

where: $E_{C\max}$ – relative elongation of makespan.

Table 5 shows the results of the simulation under the SPT dispatching rule. For each simulation, the obtained stability indicators confirmed the effectiveness and applicability of the proposed algorithm. Both the values of elongation of completion time of all jobs, $\Delta C_{\max}$, and the relative elongation of makespan, $E_{C\max}$, showed that the schedule accounting for potential technological machine failure indicates a more feasible completion time of all jobs.

For other simulated conditions, the applicability of the solutions proposed in this publication was also confirmed, as validated by mean of the performance indicators from individual simulations listed in Table 6.

The obtained values clearly indicate that the schedule incorporating service buffers gives a more feasible completion time of all jobs.

Table 5. Stability indicators – order of jobs according to the SPT rule

Sim. No.	**Executed schedule (simulation) $C'_{\max}$ [hours]**	**Elongation and relative elongation of completion times of all jobs**					
		nominal schedule			robust schedule		
		$C_{\max}$ [hours]	$\Delta C_{\max}$ [hours]	$E_{C\max}$ [–]	$C_{\max}$ [hours]	$\Delta C_{\max}$ [hours]	$E_{C\max}$ [–]
1	**56.10**		-7.18	0.87		-0.35	0.99
2	**53.88**		-4.96	0.91		1.87	1.03
3	**54.09**		-5.17	0.90		1.66	1.03
4	**56.91**		-7.99	0.86		-1.16	0.98
5	**52.60**		-3.68	0.93		3.15	1.06
6	**55.50**		-6.58	0.88		0.25	1.00
7	**56.43**		-7.51	0.87		-0.68	0.99
8	**55.88**		-6.96	0.88		-0.13	1.00
9	**53.48**		-4.56	0.91		2.27	1.04
10	**54.04**		-5.12	0.91		1.71	1.03
11	**58.31**		-9.39	0.84		-2.56	0.96
12	**52.97**		-4.05	0.92		2.78	1.05
13	**54.20**	48.92	-5.28	0.90	55.75	1.55	1.03
14	**55.33**		-6.41	0.88		0.42	1.01
15	**55.98**		-7.06	0.87		-0.23	1.00
16	**56.01**		-7.09	0.87		-0.26	1.00
17	**53.53**		-4.61	0.91		2.22	1.04
18	**56.51**		-7.59	0.87		-0.76	0.99
19	**55.18**		-6.26	0.89		0.57	1.01
20	**56.49**		-7.57	0.87		-0.74	0.99
21	**52.37**		-3.45	0.93		3.38	1.06
22	**57.52**		-8.60	0.85		-1.77	0.97
23	**54.86**		-5.94	0.89		0.89	1.02
24	**55.04**		-6.12	0.89		0.71	1.01
25	**54.83**		-5.91	0.89		0.92	1.02

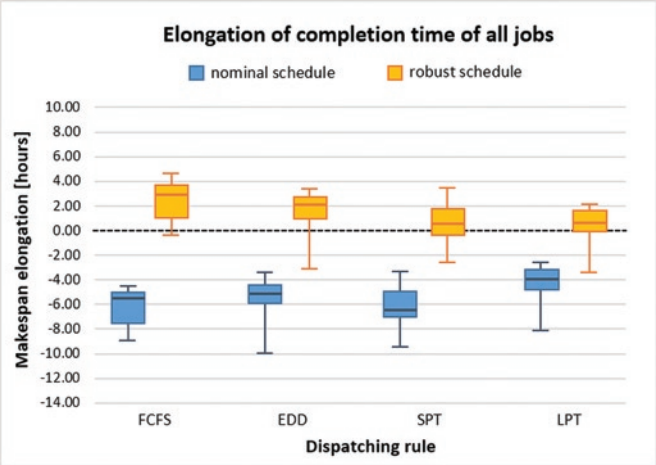


Fig. 14. Makespan elongation ΔC_{max}

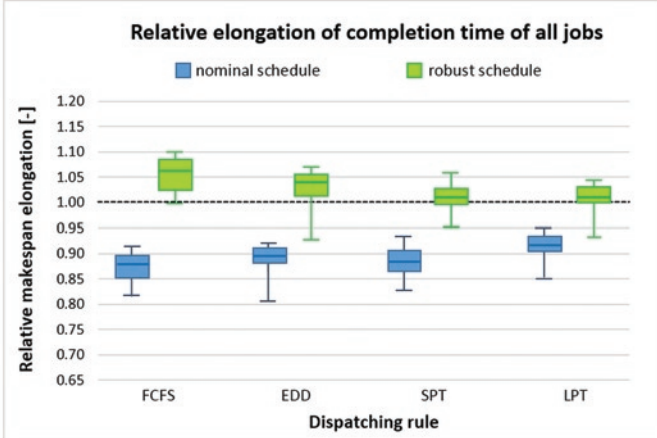


Fig. 15. Relative makespan elongation E_{Cmax}

Table 6. Mean values of the considered performance indicators

Priority rule	Executed schedule (simulation)	Elongation and relative elongation of completion times of all jobs					
		nominal schedule			robust schedule		
	$\bar{C}'_{\max}$ [hours]	$C_{\max}$ [hours]	$\Delta\bar{C}_{\max}$ [hours]	$\bar{E}_{C_{\max}}$ [–]	$C_{\max}$ [hours]	$\Delta\bar{C}_{\max}$ [hours]	$\bar{E}_{C_{\max}}$ [–]
FCFS	49.87	43.68	-6.19	0.88	52.44	2.57	1.05
EDD	47.90	42.59	-5.31	0.89	49.42	1.52	1.03
SPT	55.12	48.92	-6.20	0.89	55.75	0.63	1.01
LPT	53.14	49.10	-4.04	0.92	53.69	0.55	1.01

Figures 14 and 15 summarise the obtained values of the considered indicators, which further confirm the applicability of the proposed algorithm.

From the results of the verification and analytical works, it can be seen that the algorithm under scrutiny indicates a more feasible production completion time in the conditions allowing for the risk of technological machinery failure. This is evidenced, for instance, by the fact that for the robust schedule, the $E_{C\max}$ indicator values are close to 1, while the value of the indicator $\Delta C_{\max}$, is approximate to 0, which means that the makespans of production in the robust schedules are consistent with those obtained as a result of production simulation.

4. Summary and conclusions

Machine failure prediction has been widely investigated in numerous scientific studies. Various approaches have been proposed for the determination of information regarding the failure of technological machines. Reliable and well-developed preventive maintenance job schedules are critical to effective maintenance, particularly in the case of Time-Based Maintenance strategies.

This paper focuses on the development of a prediction algorithm using typical historical data recorded by maintenance departments. The proposed algorithm is an alternative solution to failure prediction, whose innovation, and primary advantage, consists in the implementation of Kaplan-Meier estimation to determine the characteristics of failure occurrence in time for individual technological machines of the production system, which in turn supports TBM activities. In light of these key features of the proposed prediction tool, it becomes clear that the collection of reliable data on machine failure becomes of crucial importance; it is only the adequate historical data sample size and quality that may produce reliable and factual results.

Our algorithm responds to and represents the tendency for the growing implementation of IT tools in the work of maintenance departments. Considering its potential scope of applications, it was developed as a computer program so that it is compatible with other established solutions. The verification of the proposed algorithm allowed to determine the potential failure times of technological machines. For the considered machines determined failure times were different, which means that each of them has its own failure occurrence characteristics. That confirmed the rightness and need of the TBM strategy implementation in the technical objects maintaining. The obtained data are also extremely important in the aspect of production under uncertainty. The simulation tests carried out in the second part of the publication prove that the use of the results of the proposed algorithm in the production planning allows to obtain stability of processes and determine deadlines close to the real end time of production.

The investigation works reported in this paper confirm the effectiveness of the developed prediction algorithm and indicate the need for the preventive measures to provide information on machine failure in order to improve the stability of executed processes.

Acknowledgments

The project/research was financed from the Lublin University of Technology Project – Regional Initiative of Excellence from the funds of the Ministry of Science and Higher Education on the basis of a contract No. 030 / RID / 2018/19.

References

1. Albrice D, Branch M. A Deterioration Model for Establishing an Optimal Mix of Time-Based Maintenance (TbM) and Condition-Based Maintenance (CbM) for the Enclosure System. Fourth Building Enclosure Science & Technology Conference (BEST4), Kansas City, Missouri, April 13–15, 2015.

2. Al-Hinai N, ElMekkawy TY. Robust and Stable Flexible Job Shop Scheduling with Random Machine Breakdowns Using a Hybrid Genetic Algorithm. International Journal of Production Economics 2011; 132(2): 279–291, http://dx.doi.org/10.1016/j.ijpe.2011.04.020.
3. Antosz K, Stadnicka D. Evaluation measures of machine operation effectiveness in large enterprises: study results. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2015; 17(1): 107–117, http://dx.doi.org/10.17531/ein.2015.1.15.
4. Baptista M, Sankararaman S, de Medeiros IP, Nascimento C, Prendinger H, Henriques EMP. Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling, Computers & Industrial Engineering 2018; 115: 41–53, https://doi.org/10.1016/j.cie.2017.10.033.
5. Bartochowska D, Ferenc R. Instrumenty wsparcia utrzymania ruchu w małych i średnich przedsiębiorstwach. Zeszyty naukowe Politechniki Śląskiej 2015; 80: 21–50.
6. Bei XQ, Zhu XY, Coit DW. A risk-averse stochastic program for integrated system design and preventive maintenance planning. European Journal Of Operational Research 2019; 276(2): 536–548, http://dx.doi.org/10.1016/j.ejor.2019.01.038.
7. Bräsel H, Dornheim L, Kutz S, Mörig M, Rössling I. LiSA – A Library of Scheduling Algorithms. Magdeburg University, 2001.
8. Davenport A, Gefflot C, Beck C. Slack-based Techniques for Robust Schedules. Sixth European Conference on Planning, Toledo, Spain, September 12–14, 2001.
9. Deepu P. Robust Schedules and Disruption Management for Job Shops. Bozeman, Montana, 2008.
10. Fernandes M, Canito A, Bolon-Canedo V, Conceicao L, Praca I, Marreiros G. Data analysis and feature selection for predictive maintenance: A case-study in the metallurgic industry. International Journal Of Information Management 2019, 45: 252–262, http://dx.doi.org/10.1016/j.ijinfomgt.2018.10.006.
11. Frątczak E, Sienkiewicz U, Babiker H. Analiza historii zdarzeń – Elementy teorii, wybrane przykłady zastosowań. Oficyna Wydawnicza Szkoła Główna Handlowa w Warszawie, Warszawa 2014.
12. Gao H. Bulding Robust Schedules using Temporal Potection – An Empirical Study of Constraint Based Scheduling Under Machine Failure Uncertainty. Toronto, Ontario, 1996.
13. Gao Y, Feng Y, Zhang Z, Tan J. An optimal dynamic interval preventive maintenance scheduling for series systems. Reliability Engineering & System Safety 2015; 142: 19–30, http://dx.doi.org/10.1016/j.ress.2015.03.032.
14. Gola A. Reliability analysis of reconfigurable manufacturing structures using computer simulation methods. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21(1): 90–102, http://dx.doi.org/10.17531/ein.2019.1.11.
15. Gürel S, Körpeoḡlu E, Aktürk MS. An Anticipative Scheduling Approach with Controllable Processing Times. Computers & Operations Research 2010; 37(6): 1002–1013, http://dx.doi.org/10.1016/j.cor.2009.09.001.
16. Jasiulewicz-Kaczmarek M, Bartkowiak T. Improving the performance of a filling line based on simulation, ModTech International Conference – Modern Technologies in Industrial Engineering IV, Romania, Iasi, June 15–18, IOP Conf. Series: Materials Science and Engineering 2016; 145(042024), https://doi.org/10.1088/1757-899X/145/4/042024.
17. Jensen MT. Improving robustness and flexibility of tardiness and total flow-time job shops using robustness measures. Applied Soft Computing 2001; 1: 35–52, http://dx.doi.org/10.1016/S1568-4946(01)00005-9.
18. Jian X, Li-Ning X, Ying-Wu Ch. Robust Scheduling for Multi-Objective Flexible Job-Shop Problems with Random Machine Breakdowns. International Journal of Production Economics 2013; 141(1): 112–126, https://doi.org/10.1016/j.ijpe.2012.04.015.
19. Kalinowski K, Krenczyk D, Grabowik C. Predictive-reactive strategy for real time scheduling of manufacturing systems. Applied Mechanics and Materials 2013; 307: 470–473, https://doi.org/10.4028/www.scientific.net/AMM.307.470.
20. Kempa W, Paprocka I, Kalinowski K, Grabowik C. Estimation of reliability characteristics in a production scheduling model with failures and time-changing parameters described by Gamma and exponential distributions. Advanced Materials Research 2014; 837: 116–121.
21. Kempa W, Wosik I, Skołud B. Estimation of Reliability Characteristics in a Production Scheduling Model with Time-Changing Parameters – First Part, Theory. Management and Control of Manufacturing Processes. Lublin, 2011; 7–18.
22. Kłos S, Patalas-Maliszewska J, Trebuna P. Improving manufacturing processes using simulation methods. Applied Computer Science 2016; 12(4): 7–17.
23. Lawless J. F. Statistical Models and Methods for Lifetime Data. John Wiley & Sons, 2003.
24. Leon VJ., Wu SD., Storer RH. Robustness Measures and Robust Scheduling for Job Shops. IIE transactions 1994; 26(5): 32–43, https://doi.org/10.1080/07408179408966626.
25. Liao W, Zhang X, Jiang M. An optimization model integrated production scheduling and preventive maintenance for group production. IEEE International Conference on Industrial Engineering and Engineering Management 2016; December, 936–940, http://dx.doi.org/10.1109/IEEM.2016.7798015.
26. Loska A. Scenario modeling exploitation decision-making process in technical network systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (2): 268–278, http://dx.doi.org/10.17531/ein.2017.2.15.
27. Lü Y, Zhang Y. Reliability Modeling and Maintenance Policy Optimization for Deteriorating System Under Random Shock. Journal of Shanghai Jiaotong University (Science) 2018; 23(6): 791–797, http://dx.doi.org/10.1007/s12204-018-1985-y.
28. Mehta SV., Uzsoy RM. Predictable Scheduling of a Job Shop Subject to Breakdowns. IEEE Transactions on Robotics and Automation 1998; 14(3): 365–378, https://doi.org/10.1109/70.678447.
29. Rawat M, Lad BK., Novel approach for machine tool maintenance modelling and optimization using fleet system architecture. Computers & Industrial Engineering 2018; 126: 47–62, http://dx.doi.org/10.1016/j.cie.2018.09.006.
30. Rosmaini A, Shahrul K. An overview of time-based and condition-based maintenance in industrial application. Computers & Industrial Engineering 2012; 63(1): 135–149, http://dx.doi.org/10.1016/j.cie.2012.02.002.
31. Sabuncuoglu I, Bayõz M. Analysis of reactive scheduling problems in a job shop environment. European Journal of Operational Research 2000; 126(3): 567–586, https://doi.org/10.1016/S0377-2217(99)00311-2.
32. Skołud B., Wosik I., Immune Algorithms in Production Jobs Scheduling. Zarządzanie Przedsiębiorstwem 2008; 1: 47–48.
33. Sobaszek Ł, Gola A, Kozłowski E. Job-shop scheduling with machine breakdown prediction under completion time constraint. Annals of Computer Science and Information Systems 2018; 15: 437–440, http://dx.doi.org/10.15439/2018F83.
34. Szwedzka K, Szafer P, Wyczółkowski R. Structural analysis of factors affecting the effectiveness of complex technical systems. Proceedings

of the 30th International Business Information Management Association Conference, IBIMA 2017 – Vision 2020: Sustainable Economic development, Innovation Management, and Global Growth Volume 2017, 4096–4105.
35. Timofiejczuk A, Brodny J, Loska A. Exploitation Policy in the Aspect of Industry 4.0 Concept – Overview of Selected Research. Multidisciplinary Aspects of Production Engineering 2018; 1(1): 353–359, https://doi.org/10.2478/mape-2018-0045.
36. Vonta F. Frailty or Transformation Models in Survival Analysis and Reliability. Recent Advances In System Reliability: Signatures, Multi-State Systems And Statistical Inference 2012; 237–251, http://dx.doi.org/10.1007/978-1-4471-2207-4_17.
37. Wei-Wei C, Zhiqiang L, Ershun P. Integrated Production Scheduling and Maintenance Policy for Robustness in a Single Machine. Computers & Operations Research 2014; 47: 81–91, https://doi.org/10.1016/j.cor.2014.02.006.
38. Yang BY, Liu RN, Zio E. Remaining Useful Life Prediction Based on a Double-Convolutional Neural Network Architecture. IEEE Transactions On Industrial Electronics 2019; 66(12): 9521–9530, https://doi.org/10.1109/TIE.2019.2924605.
39. Zhang F, Shen J, Ma Y. Optimal maintenance policy considering imperfect repairs and non-constant probabilities of inspection errors. Reliability Engineering and System Safety 2020; 193: 1–12, http://dx.doi.org/10.1016/j.ress.2019.106615.
40. Zhao X, He S, He Z, Xie M. Optimal condition-based maintenance policy with delay for systems subject to competing failures under continuous monitoring. Computers & Industrial Engineering 2018; 124: 535–544, http://dx.doi.org/10.1016/j.cie.2018.08.006.

Łukasz SOBASZEK
Arkadiusz GOLA
Antoni ŚWIĆ
Department of Production Computerisation and Robotisation
Faculty of Mechanical Engineering, Lublin University of Technology
ul. Nadbystrzycka 36, 20-816 Lublin, Poland

E-mails: l.sobaszek@pollub.pl, a.gola@pollub.pl, a.swic@pollub.pl

Article citation info:
LI Y, WANG K. Modified convolutional neural network with global average pooling for intelligent fault diagnosis of industrial gearbox. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 63–72, http://dx.doi.org/10.17531/ein.2020.1.8.

Yaxin LI
Kesheng WANG

MODIFIED CONVOLUTIONAL NEURAL NETWORK WITH GLOBAL AVERAGE POOLING FOR INTELLIGENT FAULT DIAGNOSIS OF INDUSTRIAL GEARBOX

DIAGNOSTYKA BŁĘDÓW PRZEKŁADNI PRZEMYSŁOWYCH Z WYKORZYSTANIEM ZMODYFIKOWANEJ SPLOTOWEJ SIECI NEURONOWEJ Z GLOBALNYM UŚREDNIENIEM WARTOŚCI DLA POSZCZEGÓLNYCH KANAŁÓW

Gearboxes are key transmission components and widely used in various industrial applications. Due to the possible operational conditions, such as varying rotational speeds, long period of heavy loads, etc., gearboxes may easily be prone to failure. Condition Monitoring (CM) has been proved to be an effective methodology to improve the safety and reliability of gearboxes. Deep learning approaches, nowadays, further enable the CM with more powerful capability to exploit faulty information from massive data and make intelligently diagnostic decisions. However, for most of conventional deep learning models, such as Convolutional Neural Network (CNN), a large amount of labelled training data is a prerequisite, while to obtain the labelled data is usually a laborious and time-consuming job and sometimes even unattainable. In this paper, to handle the case of only a limited labelled data is available, a modified convolutional neural network (MCNN) is proposed by integrating global average pooling (GAP) to reduce the number of trainable parameters and simplify the architecture of deep learning model. The proposed MCNN improves the traditional CNN's ability in fault diagnosis with limited labelled data. Two experimental gearbox datasets are utilized to demonstrate the effectiveness of the proposed MCNN method. Compared with traditional deep learning approaches, namely LSTM, CNN and its variant methods, the experimental results show that the proposed MCNN with higher discrimination and generalization ability in fault classification and diagnostics under the scenario of limited labelled training samples.

Keywords: *modified convolutional neural network, global average pooling, intelligent fault diagnosis, industrial Gearbox.*

Przekładnie stanowią kluczowe elementy układów napędowych i jako takie znajdują szerokie zastosowane w przemyśle. Ze względu na warunki eksploatacji, takie jak różne prędkości obrotowe czy długie okresy pracy pod dużym obciążeniem itp., przekładnie mogą łatwo ulegać uszkodzeniom. Udowodniono, że monitorowanie stanu skutecznie poprawia bezpieczeństwo i niezawodność przekładni. Podejścia oparte na uczeniu głębokim umożliwiają ponadto monitorowanie stanu z większym wykorzystaniem informacji o błędach pochodzących z dużych zbiorów danych i podejmowanie inteligentnych decyzji diagnostycznych. Jednak w przypadku większości konwencjonalnych modeli uczenia głębokiego, takich jak splotowe sieci neuronowe (convolutional neural networks, CNN), wymagana jest duża ilość etykietowanych danych uczących, których pozyskanie jest zwykle zadaniem praco- i czasochłonnym, a czasem wręcz niemożliwym do wykonania. W niniejszej pracy, przedstawiono zmodyfikowaną splotową sieć neuronową (modified convolutional neural network, MCNN), która rozwiązuje problem dostępności danych etykietowanych poprzez zastosowanie globalnego uśrednienia względem kanałów (global average pooling), co pozwala na zmniejszenie liczby możliwych do wyuczenia parametrów i uproszczenie architektury modelu głębokiego uczenia. W porównaniu do tradycyjnych sieci CNN, proponowana sieć MCNN zwiększa możliwości diagnozowania błędów przy ograniczonych danych etykietowanych. Skuteczność proponowanej metody wykazano na przykładzie dwóch zbiorów danych doświadczalnych dotyczących błędów przekładni. Wyniki eksperymentalne pokazują, że, w porównaniu z tradycyjnymi metodami uczenia głębokiego, takimi jak LSTM, CNN oraz warianty tej ostatniej, proponowane podejście MCNN daje większe możliwości rozróżniania i uogólniania podczas klasyfikacji i diagnostyki błędów w przypadku ograniczonej dostępności etykietowanych danych uczących.

Słowa kluczowe: *zmodyfikowana splotowa sieć neuronowa, globalne uśrednienie względem kanałów, inteligentna diagnostyka błędów, przekładnia przemysłowa.*

1. Introduction

High transmission ratio, strong load-bearing and high efficiency makes modern gearboxes are always considered to be critical components in various industrial applications, such as wind turbine generator system, helicopter main speed reducer, aerospace engineering and etc [27, 38]. In real practice, however, gearboxes will inevitably be subjected with dynamic heavy-duty loads under complex operating conditions, making the breakdown or even accidents of the engineering system [3, 4, 8, 32, 39]. Therefore, it is of great significance to develop the condition monitoring and fault diagnostic techniques for the gearboxes.

Most of the modern gearbox fault diagnostic methods utilize vibration analysis to extract the fault features, and then make decision according to sophisticated signal processing techniques or expert knowledge of diagnosticians [1, 2, 6, 9, 18, 33]. For instance, Feng et al. [10] successfully introduced the Vold-Kalman filter into time-frequency analysis to extract fault features of the planetary gearbox under unstable operation conditions. Tang et al. [28] firstly presented a novel fault detection method to identify the categories of gearbox

failures on the strength of hierarchical instantaneous energy density dispersion entropy (HIEDDE) and dynamic time warping (DTW). However, for these approaches based on vibration analysis, large amounts of signal processing efforts and abundant expert diagnostic experience are generally required to extract and analyze fault characteristics from the measured vibrations.

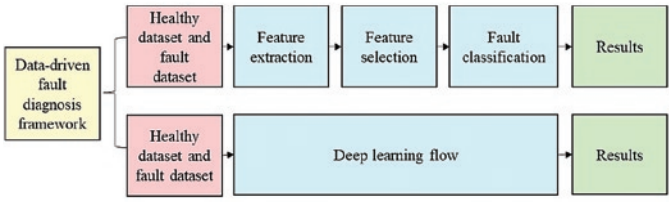


Fig. 1. Data-driven fault diagnosis framework.

Thanks to the recent advanced progress made by artificial intelligence and machine learning techniques, intelligent fault diagnosis receives an increasing research attention in the field of condition monitoring and fault diagnosis [12, 13, 25, 39, 43]. Methods, such as artificial neural network (ANN), back-propagation neural network (BPNN) and support vector machine (SVM) have become research focus. For instance, Tyagi et al. [29] successfully constructed a hybrid artificial neural network (ANN) classifier for gearbox diagnosis. The hybrid classifier consists of data preprocessing with discrete wavelet transform (DWT), genetic algorithm (GA) and back-propagation neural network (BPNN). Zhang et al. [42] developed a multivariable ensemble-based incremental support vector machine (MEISVM) and applied it into the compound failure detecting of roller bearings. Despite their successes, the outstanding performance of the intelligent diagnostic methods alike heavily count on the accuracy of the manually extracted and selected features. This typical route of intelligent method is shown in the upper half in Fig. 1. For this route, advanced signal processing techniques are usually required for data pre-processing. Moreover, for the shallow learning model (such as BPNN, SVM) employed for gearbox fault diagnosis, the diagnostic ability of the model is relied heavily on the quality of the extracted fault features. Unfortunately, for the cases of large amount of measured industrial data with unguaranteed data quality, its diagnostic capability will naturally exhibit insufficient with the increase of the data amount.

To tackle these issues, deep learning route shown in the lower half in Fig. 1 with the structure of deep learning model of multi-layer nonlinear modelling solutions, comes into the recent research focuses and provides a straightforward end-to-end learning process from the measured input signal to the output diagnostic results, which completely eliminate the challenges of manual feature extractions and selections [13, 44]. Deep learning methods utilize the deep architectures to constitute hierarchical feature representations to discover the distributed feature representations of data. Due to its powerful capability of perception, self-learning, modelling and characterization, recent years have witnessed the tremendous progress of deep learning techniques in various fields, mainly including image processing, speech recognition and fault diagnosis. For examples to fault diagnosis, Yu et al. [37] exploited stacked denoising auto-encoder (SDAE) and gated recurrent unit neural network (GRUNN) to enhance the capability of anti-noise and the adaptive ability to time-varying rotational speed during the fault diagnosis of the planetary gearbox. Liu et al. [22] automatically extracted features from vibration signal by combining batch normalization with deep belief network (DBN), which achieves more precise performance than DBN and other conventional approaches in wind turbine gearbox diagnosis. Furthermore, Other deep neural network such as deep residual network (DRN) [40], recurrent neural network (RNN) [21], generative adversarial networks (GAN) [26], long short-term memory networks (LSTM) [19, 36] and especially convolutional neural networks (CNN) are seriously and widely investigated in gearbox fault diagnosis. Specifically, to CNN, Chen et al. [7] applied CNN to adaptively learn fault features and classify fault patterns with extreme learning machine (ELM) for mechanical faults. Jiao et al. [17] developed a deep coupled dense convolutional network (CDCN) to diagnose the faults of planetary gearbox, which could relieve gradient vanishing in deep architecture and realize two-stage information fusion.

Even though various successful cases on the applications of CNN in fault diagnosis have been reported, the works generally employed massive labelled measured samples to train a deep network. Nevertheless, it is difficult to acquire an enough number of labelled samples in real industrial application, especially for certain faulty scenarios which seldomly occur. Recently, transfer learning may be promising in this problem, and some works about transfer learning-based fault diagnosis have been reported [15]. The unsupervised domain adaptation is the major branch of this framework [5, 14, 34]. By adapting the feature distribution between two domains, the diagnostic model can generalize well to the unseen conditions where no labelled data can be used for model training. Although these methods avoid to use labelled data in target domain, a large amount of unlabelled data (similar to the data in source domain) are generally necessary. In addition, due to the complicated and deep multi-layer structures, the parameter optimization of CNN model may lead to a huge computational burden. In this scenario, promoting the precision of diagnosis, accelerating the training speed of CNN as well as boosting the generalization ability and robustness under the small number of labelled samples become a critical issue to research. To this end, in this paper, a modified convolutional neural network (MCNN) for the fault diagnosis was proposed in which the global average pooling (GAP) is introduced into the internal structure of CNN model to replace the traditional fully connected layer where the majority of parameters for training are contained. By doing so, compared with traditional CNN, it enables the MCNN an improved capability to deal with fault classification problem with limited labelled samples. The MCNN is validated by two gearbox datasets from PHM 2009 conference data challenge and measured experimental data at University of Electronic Science and Technology of China. The experimental results demonstrate the advantages of the proposed MCNN method in accuracy of fault classification, time-saving of training model, more important, the method is more effective for fault diagnosis under condition of limited number of labelled samples.

The structure of the paper is arranged as follows. In section 2, the basic theory of CNN is briefly described. The method of MCNN is introduced in section 3. In section 4, the two experimental application of proposed method is analysed. Finally, the summary is concluded in section 5.

2. Methods

2.1. Traditional architecture of CNN

As one of the most representative deep learning algorithms, CNN is a combination of convolutional computation and deep structure, which is generally composed of three parts, i.e., input layer, the feature descriptor and the classifier. The feature descriptor consists of multiple convolution layers, activation layers, pooling layers. The input signal is mapped to the feature space of the CNN hidden layers to extract the features of the input data. The classifier is composed of one or several fully connected layers, namely a multi-layer perceptron classifier, for fusion and classification of the extracted features. The input layer of CNN is to pre-process multidimensional data, usually referring to one-dimensional data, two-dimensional data, or three-dimensional data. As the core of CNN, the convolution layer, containing multiple convolution kernels, is to perform feature learning and

extraction from the input signal. Each convolution kernel corresponds to a weight matrix and a bias vector, similar to a neuron of a feed-forward neural network. The convolution kernels sweep through the input features with a pre-set stride, and obtain the activated feature maps in the receptive field.

Supposing that the input signal, and the filer $w \in R^n$, the convolution process can be depicted as:

$$X_i^{(k)} = \sum_{c=1}^{C} W_i^{(c,k)} * X_{i-1}^{(c)} + B_i^{(k)} \tag{1}$$

where i represents the index of convolutional layer, k represents the index of feature map in i^{th} layer, c means the number of convolutional kernel in i^{th} layer, $*$ means the convolution operation, $X_{i-1}^{(c)}$ is the input feature map of $(i-1)^{th}$ layer, $X_i^{(k)}$ is the output feature map, $W_i^{(c,k)}$ and $B_i^{(k)}$ donates the weights and bias of the convolution kernels respectively.

By introducing nonlinear activation function into the network model, the ability of feature representation will be further enhanced. The generally utilized activation functions include sigmoid, tanh, rectified linear units (ReLU), etc. These functions are listed in equation (2). Among them, ReLU is one of the most noteworthy functions with efficient gradient descent ability and avoiding gradient explosion and disappearance during the training process:

$$\begin{cases} sigmoid: f = \dfrac{1}{1+e^{-y}} \\ tanh: f = \dfrac{e^{y}-e^{-y}}{e^{y}+e^{-y}} \\ ReLu: f = max(0, y) \end{cases} \tag{2}$$

After the convolutional operations, the output feature maps are delivered to the pooling layer for down-sampling. The widely used max-pooling is to divide the feature maps into a series of blocks without overlapping and extract the maximum value in each block as the eigenvalue of the window while discarding other points. The max-pooling process can be defined as equation (3):

$$X_{i+1}^{(k)} = max \underset{(i-1)l+1 \le t \le il}{max} X_i^{(k)}(t) \tag{3}$$

where $X_i^{(k)}(t)$ represents the feature map after convolution operation of the kth neuron at the ith layer, l denotes the width of a local area for max-pooling, $X_{i+1}^{(k)}(i)$ is the output feature map after max-pooling.

The fully connected layers are located at the last part of CNN with the purpose of nonlinearly combining the extracted features and mapping them into output labels. The softmax function is used at the final output layer to calculate the probability distribution for each label, whose mathematical expression can be described in equation (4):

$$\mathrm{H}_{,\left(\mathrm{x}^{(i)}\right)} = \begin{bmatrix} p\left(\mathrm{y}^{(i)} = 1|\mathrm{x}^{(i)};,\right) \\ p\left(\mathrm{y}^{(i)} = 2|\mathrm{x}^{(i)};,\right) \\ \cdots \\ p\left(\mathrm{y}^{(i)} = \mathrm{m}|\mathrm{x}^{(i)};,\right) \end{bmatrix} = \frac{1}{\sum_{\mathrm{j}=1}^{\mathrm{k}} \exp\left(,_{\mathrm{j}}^{\mathrm{T}} \mathrm{x}^{(i)}\right)} \begin{bmatrix} \exp(,_{1}^{\mathrm{T}} \mathrm{x}^{(i)}) \\ \exp(,_{2}^{\mathrm{T}} \mathrm{x}^{(i)}) \\ \cdots \\ \exp(,_{\mathrm{k}}^{\mathrm{T}} \mathrm{x}^{(i)}) \end{bmatrix} \tag{4}$$

where the interregional of $y^{(i)}$ is $\{1,2,\ldots m\}$, m is the number of classifications, and θ is the assemblage of the arguments of the model.

2.2. A discussion on the shortcoming of the CNN

LeNet, as the pioneering and widely used architecture of CNN, basically established by convolutional layers, pooling layers and fully-connected layers. Most of researches for fault diagnosis utilized the simplified and improved LeNet-5 (containing 5 convolutional layers). For CNN with LetNet architecture, though it has been widely acknowledged, a large amount of training data for a proper modelling is a prerequisite and this may be largely attributed to the architecture of the LeNet with numerous network parameters. Specifically, in this architecture, the last set of feature maps are flattened into one-dimensional feature vector, and each feature is connected to each neuro in the first fully-connected layer. In this manner, the extracted features will be mapped into label space. However, it should be noted that, even at the end of feature maps, considerable amount of network parameters still exists and need to be trained. As a result, it makes the proper training of the model with small of amount of data becomes a tough issue. To give an intuitive presentation and quantitative analysis, three famous CNN architectures in computer vision, i.e., LeNet-5, AlexNet and VGG-16, as examples, the distributions of parameters between front convolutional block and later fully-connected layers are shown in Table 1. It is clear that the vast majority of parameters are distributed in the fully-connected layers. Consequently, to remove or modify the complex part of the model by reducing the number of trainable parameters within the CNN structure, at the same time as large as possible to remain the feature extraction representation results, can be a promising solution to improve the capability of the model, especially enabling the model to deal with small amount of labelled data samples which will be beneficial to the real engineering practice.

Table 1. The distributions of parameters in CNN with three typical architectures

Architecture	Parameters distribution (%)	
	Convolutional block	Fully-connected layer
LeNet-5	2.8	97.2
AlexNet	3.8	96.2
VGG-16	10.6	89.4

2.3. Global average pooling (GAP)

The novel global average pooling (GAP) [20] is, therefore, introduced in this section with which fully connected layers of CNN is superseded. For each feature map at the end of pooling layers, we take the average value of each feature vector directly maps to a category label or an output node. This process was called global average pooling. The original fully-connected layers are replaced. By doing so, a tremendously reduction of the number of parameters needed to be trained is realized and the computational burden of training the model is decreased. More important, though it simplified the CNN model, it still completely remains the key convolutional layers and therefore the ability of feature representation still remained. Furthermore, it gives an extra capability to the model to deal with the training problems with small amount of data samples. The detailed illustrations of the traditional fully-connected layers and global average pooling layer are shown in Fig. 2.

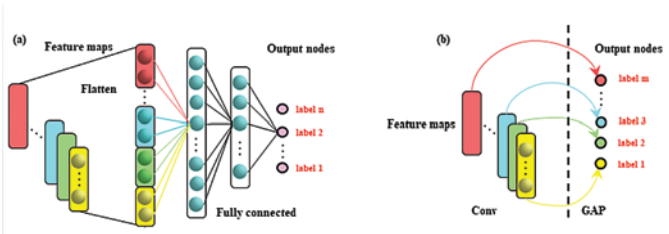


Fig. 2. An illustration of fully connected layer and global average pooling layer. (a) fully-connected layer; (b) global average pooling layer

2.4. Modified CNN

The structure of MCNN consists of input layers, convolutional layers, max-pooling layers, dropout layers, global average pooling layers. The deep network has a total of 19 layers. We utilize 5 convolutional layers to learn features from raw data (follow the architecture of LeNet-5). The ReLU function is employed as the activation function. A max-pooling layer is performed after per convolution operation. Since the number of last sets of feature maps is equal to the category labels or output nodes where the global average operation will be performed. Thus, an additional convolutional layer, called task specific layer, is added to revise the number of the out-feature maps. The number of convolutional kernels in the task specific layer equals to the category labels. The detailed structure and parameter setting of MCNN constructed in this paper are illustrated in Fig. 3. Referring to the literature [41], larger values, such as 128 and 64, are selected as the kernel sizes for the front two convolutional layers to capture essential features and reduce high frequency noise for 1D vibration signal. With the increase of network depth, the number of kernels also increase from 16 to 256, which helps to learn hierarchical feature representation. It worth noting that, as a regularization approach, dropout layers are integrated after each specific layer to prevent overfitting.

In the following comparative analysis, the same construction and parameters in the feature descriptor are chosen in the traditional CNN in order to conduct a fair comparison. The following flatten layer is employed for transforming the high dimensional feature maps to one-dimension feature vector. Afterwards, 3 fully connected layers are established to integrate local information with class discrimination in convolutional blocks and map the learned distributed feature representation to the sample label space. The detailed CNN and MCNN structures can be seen in Fig. 3.

2.5. The intelligent fault diagnosis framework with MCNN

In this paper, an intelligent fault diagnosis framework for gearbox with MCNN is proposed and listed in figure 4. There are 3 steps in total: (1) data processing, (2) train the MCNN model, and (3) fault diagnosis of gearbox.

(a) Data processing

Measured vibration signals are collected from accelerators and directly input into the fault diagnostic framework without any manual feature extraction and selection. In this way, an end-to-end fault diagnostic framework is realized and the loss of data information caused by human interventions or advanced signal processing techniques are minimized. According to the generally used sample size of deep learning researches [30, 41], the raw signal is partitioned into a series of fixed-length segments by shifting the window with a constant stride, and then the training and testing data sets are selected from the whole measured vibrations without repetition.

(b) MCNN Model learning

Model training consists of two stages: forward calculation and loss backward propagation. In the forward calculation stage, training samples are fed into the MCNN model and predicted outputs. Then, the loss between the predicted outputs and the real outputs are backward propagated to optimize the network parameters layer by layer. The method of the optimizer utilized is stochastic gradient descent (SGD) technique. The optimization process can be represented as:

$$\theta = \theta - \cdot * \nabla_{,} J\left(\theta; x^{(i)}; y^{(i)}\right) \tag{5}$$

where θ is the collection of network parameters, $x^{(i)}$ and $y^{(i)}$ represent the input sample and corresponding label, $J(\cdot)$ is the loss function, η is the learning rate and ∇_θ denotes the gradients.

(c) Fault diagnosis of gearbox

After completing the training of the MCNN model, the diagnostic model is deployed for fault classification. And the testing samples are fed into the model for validation. The diagnostic accuracy of model is defined to evaluate the performance of the network.

$$Accuracy = \frac{\left|x : x \in D \wedge \hat{y}(x) = y(x)\right|}{\left|x : x \in D\right|} \tag{6}$$

where D is the set of test data, x is the input sample, $y(x)$ is the truth label of x, $\hat{y}(x)$ is the label predicted by the diagnosis model.

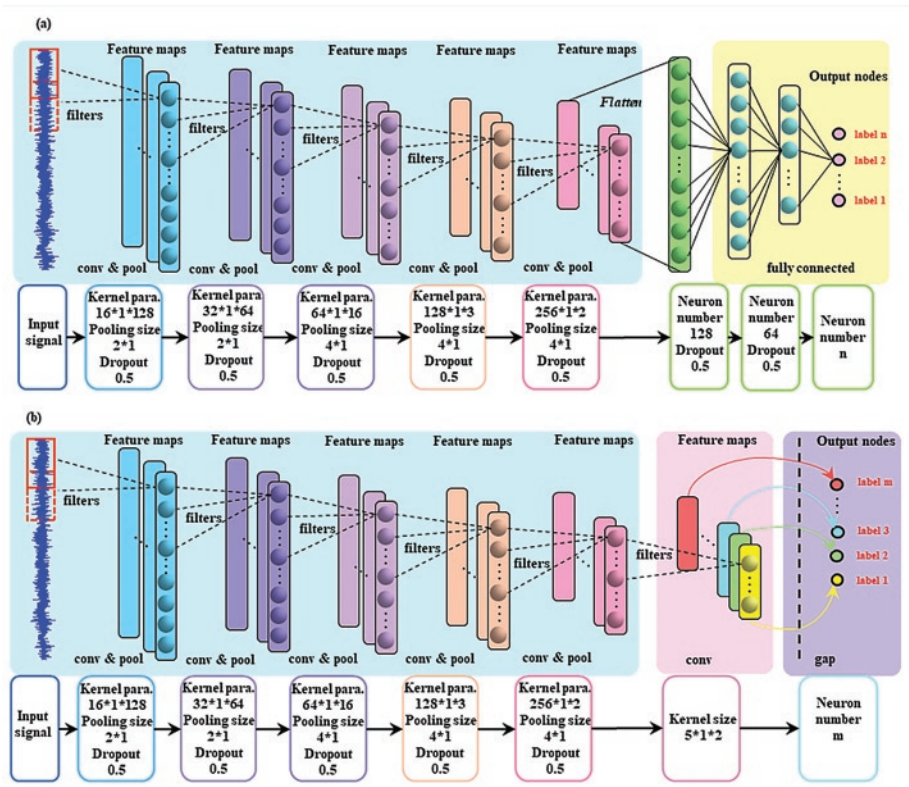


Fig. 3. An illustration of CNN and MCNN: (a) CNN; (b) MCNN

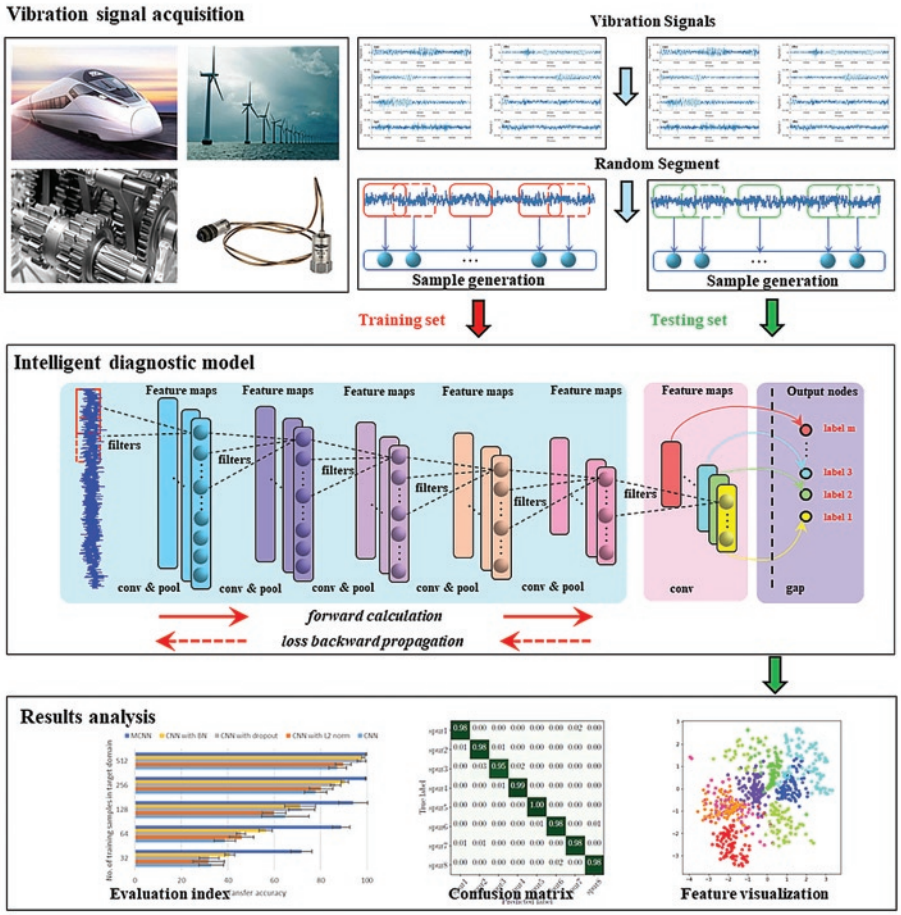


Fig. 4. The intelligent fault diagnosis framework with MCNN

3. Experimental studies

3.1. Description of dataset from 2009 challenge dataset of the Prognostics and Health Management (PHM) society

The 2009 challenge dataset from Prognostics and Health Management (PHM) society is first used to validate the proposed MCNN [23]. The datasets are measured from the gearbox shown in Fig.5. Fig. 5 (a) illustrates the constitution of the fixed shaft gearbox and the position of accelerometers and tachometer. The gearbox consists of 3 shafts, 4 spur gears and 6 bearings, as is shown in Fig. 5 (b). Vibration signals of the spur gear with 8 health conditions, 5 shaft speed conditions including 30, 35, 40, 45 and 50 Hz, under the same amount of high load are collected. The Table 2 lists the specific 8 failure modes of the gearbox. These experiments can fundamentally cover the frequently occurred faults in gearbox. These faults are artificially introduced to machines so as to simulate diverse health conditions. The sampling frequency is 66.67 kHz. There are 533312 points under each fault mode and operation condition, hence 6144 data points with a 4096 stride are collected for one sample to guarantee that there is abundant fault information for each sample. There are 5208 samples in total. The time waveforms for each health condition are shown in

Table 2. Descriptions of detailed fault patterns

Label	Gear				Bearing						Shaft	
	32T	48T	80T	96T	IS:IS	ID:IS	OS:IS	IS:OS	ID:OS	OS:OS	Input	Output
1	G	G	G	G	G	G	G	G	G	G	G	G
2	C	G	E	G	G	G	G	G	G	G	G	G
3	G	G	E	G	G	G	G	G	G	G	G	G
4	G	G	E	Br	B	G	G	G	G	G	G	G
5	C	G	E	Br	In	B	O	G	G	G	G	G
6	G	G	G	Br	In	B	O	G	G	G	Im	G
7	G	G	G	G	In	G	G	G	G	G	G	Ks
8	G	G	G	G	G	B	O	G	G	G	Im	G
IS = input shaft; :IS = input side; ID = idler shaft; OS = output shaft; :OS= output side. G: good; C: chipped; E: eccentric; Br: broken; B: ball; In: innerrace; O: outer race; Im: imbalance; Ks: keyway sheared.												

Fig. 6. The training and testing data set are randomly selected from the whole dataset without repetition. The number of samples in the training is selected as 128, 256, 512, 1024, and 4096 successively, while the number of testing sets is 1000.

3.2. Description of data from the Drivetrain Diagnostics Simulator (DDS) test rig at UESTC

The second dataset is from the Drivetrain Diagnostics Simulator (DDS) test rig at University of Electronic Science and Technology of China (UESTC). The layout of the test rig is shown in Fig. 7. The accelerometer is mounted on the one-stage planetary gearbox for the collection of vibration signals. The structure of the one-stage planetary gearbox is shown in Fig. 8 (a), which is constituted by a sun gear, 4 planet gear, planet carrier and ring gear. Four different kinds of faults in the sun gear of the one-stage of the planetary gearbox is shown in Fig. 8, including tooth wear, tooth broken, tooth missing and root crack. For each sun gear health condition, 6.39 seconds of data is collected under 2 different loads (0A, 1.3A) and 3 different rotational speeds (30Hz, 40Hz and 50Hz), with the sampling frequency of 30.72 kHz. 196,608 points are collected for each health condition under each operation condition, and 2048 data points with a 1000 stride are cut for each data sample. Fig. 9 shows the typical original vibration waveforms for each health condition.

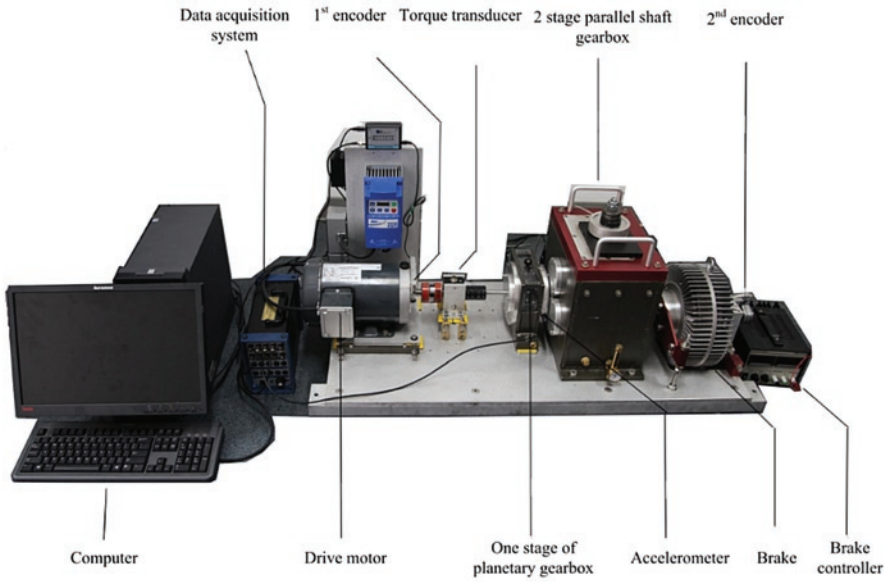


Fig. 7. DDS test rig

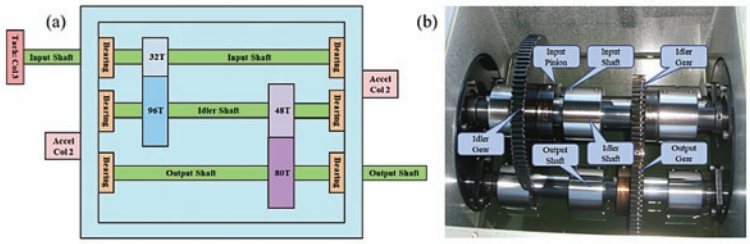


Fig. 5. The gearbox in the 2009 Challenge Data of PHM society: (a) Schematic diagram; (b) Overview of the gearbox

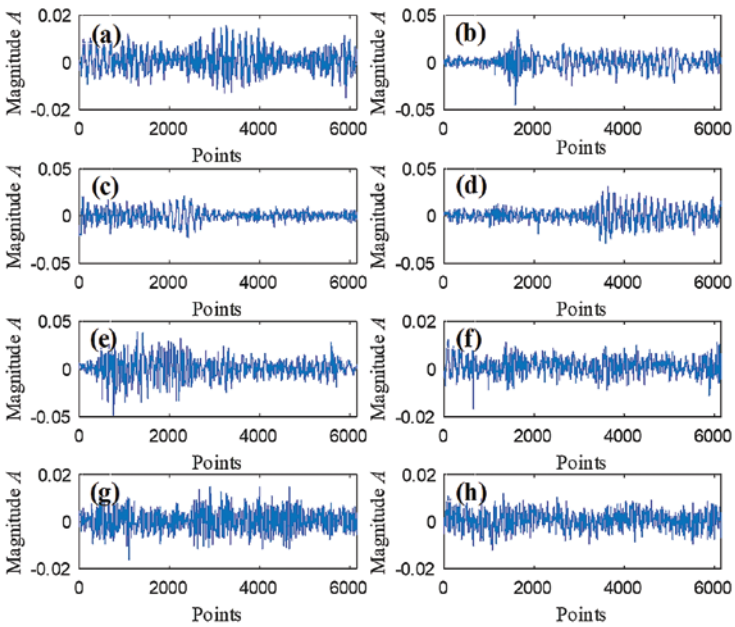


Fig. 6. Collected vibration signals of 8 machine conditions: (a) spur1 (b) spur2 (c) spur3 (d) spur3 (e) spur5 (f) spur6 (g) spur7 (h) spur8.

Therefore, 49 samples are generated, and two sets of samples with different sizes are randomly selected, one of which is set as the training set while another is set as the testing set. We set the number of samples in the training set to 128, 256, 512, 1024, and 4096 in turn, and the sample size in the testing set is 800.

3.3. Comparative methods

The proposed MCNN will be compared with other intelligent fault diagnostic methods, including (1) support vector machine (SVM) [35], (2) random forest (RF) [11], (3) long short-term memory (LSTM) [19], (4) CNN [41], (5) CNN with l2-norm [24], (6) CNN with batch normalization (BN) [31]. among them, RF and SVM are two of the most commonly used models in machine learning. RF, containing multiple decision trees, is a classifier that uses multiple trees to train and predict samples. SVM is a nonlinear kernel classifier that categorizes the data based on supervised learning. LSTM is a time-cycle neural network. L2-norm and BN are two different regularization algorithms, which can effectively improve the performance of CNN and prevent overfitting.

The related architecture parameter settings of the other methods are listed as follows. (1) RF: the number of trees and random feature subset are separately set as 500 and $\sqrt{m}$. (2) SVM: radial basis function (RBF) is introduced as the kernel function of SVM, besides, the arguments of RBF and the penalty factor are intelligently optimized by genetic algorithm (GA). The maximum generation and the number of populations in GA is set to 50 and 20 respectively. And the searching range of parameters in SVM is set to [0, 100]. (3) LSTM: By referring to [19], the dimension of hidden layer is 128 and two RNN layers are stacked. (5) CNN: the architecture and parameter settings have been listed in 2.3. (6) CNN with l2-norm: l2-norm regularization is introduced to the parameters of CNN with a weight of 1e-2. It should be noted that the popular statistical features in time domain and frequency domain [11], such as

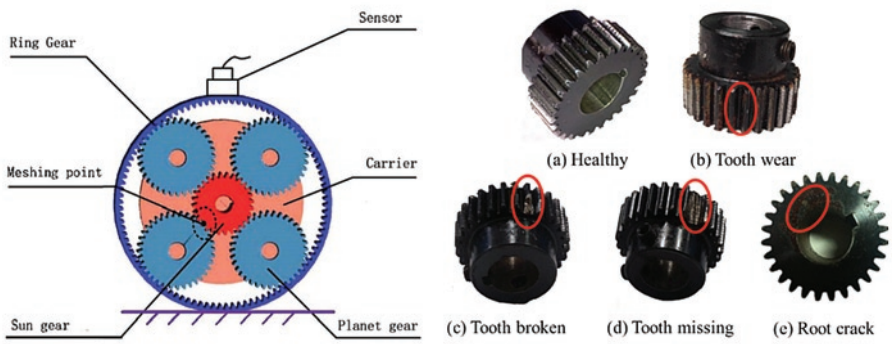


Fig. 8. The gearbox in DDS. (a) Schematic diagram; (b) 5 health conditions of sun gear

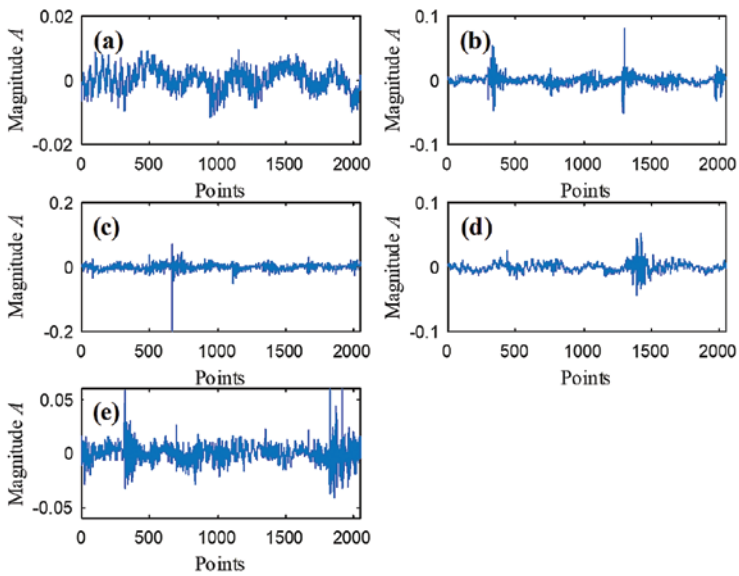
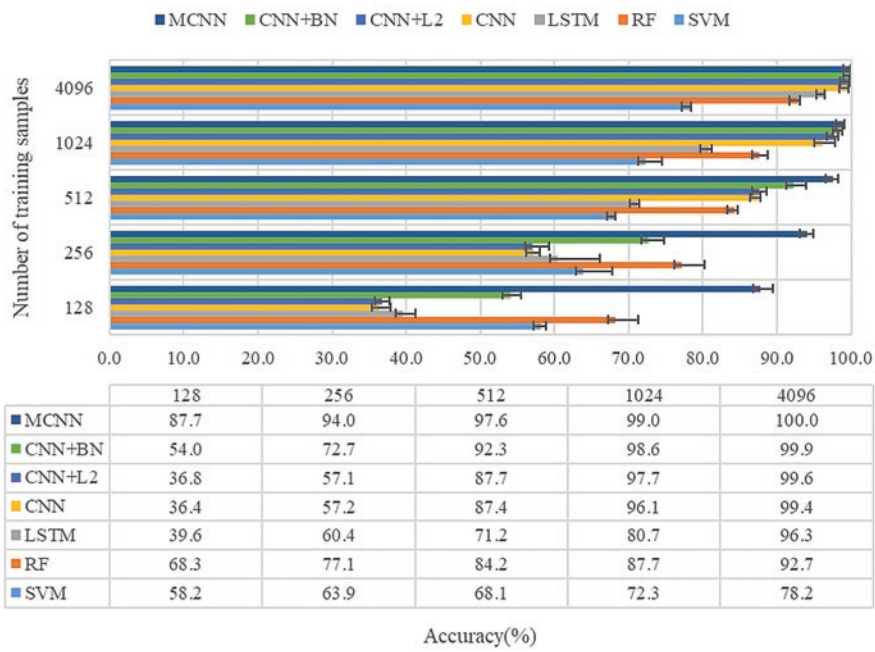


Fig. 9. Collected vibration signals of 5 machine conditions: (a) healthy(b) tooth broken(c) tooth crack (d) tooth missing (e) tooth wear.



	128	256	512	1024	4096
MCNN	87.7	94.0	97.6	99.0	100.0
CNN+BN	54.0	72.7	92.3	98.6	99.9
CNN+L2	36.8	57.1	87.7	97.7	99.6
CNN	36.4	57.2	87.4	96.1	99.4
LSTM	39.6	60.4	71.2	80.7	96.3
RF	68.3	77.1	84.2	87.7	92.7
SVM	58.2	63.9	68.1	72.3	78.2

Fig. 10. Diagnosis results of PHM09 gearbox using different sample sizes with 7 methods

root mean square, kurtosis, skewness, etc, are used as the input for shallow methods, i.e., SVM and RF, while the raw signal are sent to deep methods, i.e., LSTM, CNN and its variant methods, MCNN, due to the adaptive feature learning ability.

4. Results and discussions

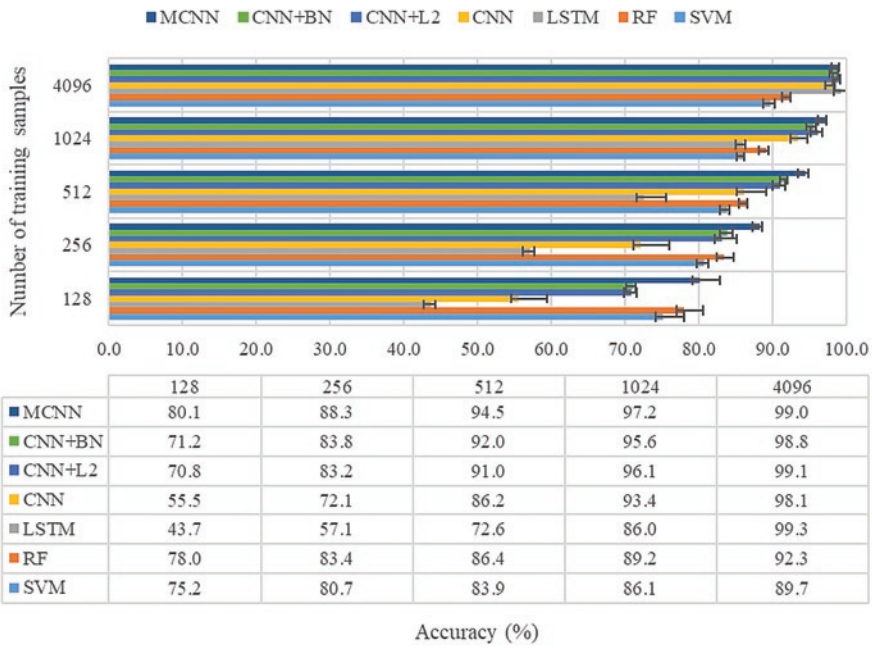
The results of two case studies versus different number of training samples for diverse methods are given in Figs. 10 and 11. Each result is an average of 10 random repeats, the average value and variance of classification accuracy of testing samples are shown in figures.

Deep and shallow learning structure comparison: As we can see in the Fig. 10, when sufficient training samples are provided, diagnostic models based on deep learning perform superior to the shallow machine learning methods, and all the intelligent diagnostic methods have achieved good classification results and it indicates that deep network structure has stronger feature learning ability than shallow network architecture.

Training sample size comparison: When the size of the samples decreases, the performance of traditional deep learning approaches presents a dramatic decline. CNN shows a worst sharp decrease in classification accuracy and it reveals that CNN is prone to over-fitting and has a weak generalization ability with small training sample size. BN and L2-norm are two frequently-used algorithms to prevent CNN from over-fitting. As is shown in the chart of Fig. 10, BN with CNN indeed has some improvements, compared to basic CNN, under small sample conditions.

Comparisons with the proposed MCNN: As is shown in Fig. 10 and Fig. 11, MCNN exhibits a significant improvement in terms of fault classification accuracy compared with other models. Compared with CNN, CNN with L2-norm, and CNN with BN, MCNN increases classification accuracy with 51.3%, 50.9%, 33.7% respectively when using the number of training samples with 128. This remarkable improvement indicates that, with introduction of GAP into the network structure, the proposed MCNN exhibits superior advantages in feature learning and generalization with small number of training samples. It enables the proposed MCNN method can be a promising tool to deal with the real-world challenge for fault diagnosis with only limited labelled data available.

It also should be noted that classical machine learning algorithms, such as RF and



	128	256	512	1024	4096
MCNN	80.1	88.3	94.5	97.2	99.0
CNN+BN	71.2	83.8	92.0	95.6	98.8
CNN+L2	70.8	83.2	91.0	96.1	99.1
CNN	55.5	72.1	86.2	93.4	98.1
LSTM	43.7	57.1	72.6	86.0	99.3
RF	78.0	83.4	86.4	89.2	92.3
SVM	75.2	80.7	83.9	86.1	89.7

Fig. 11. Diagnosis results of DDS planetary gearbox using different sample sizes with 7 methods

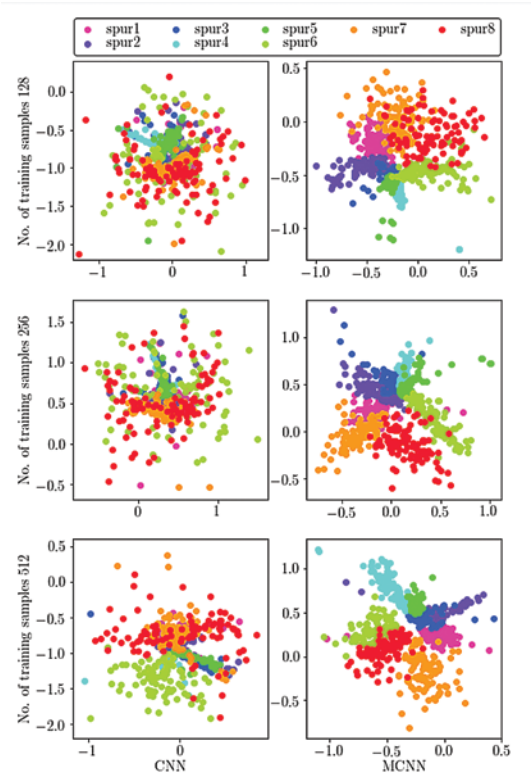


Fig. 12. Features visualization in PHM2009 dataset with CNN and MCNN

SVM, are shallow structure methods and are capable of handling small sample size. From the Fig. 10, RF and SVM are also performed better than most of the deep learning methods, such as LSTM, CNN and its variant methods, when the number of training samples are 128 and 256 respectively. Nevertheless, the proposed MCNN, though with the limited training samples of 128 and 256, it still achieves a higher diagnostic accuracy than RF and SVM. The traditional shallow methods of RF and SVM rely more on the quality of the artificial feature extraction. The classical statistical features in time-domain and frequency-domain are utilized in this paper [11]. These features have a certain sensitivity and resolving power for different faults, whereas, they possess distinct fault description capabilities for different application objectives. However, MCNN is a framework based on deep learning with remarkable adaptive feature learning capabilities. Experimental results illustrate that MCNN still achieves superior performance even under small sample conditions. Similar conclusions can also be demonstrated in Fig. 11.

In order to visually verify the effectiveness of the MCNN algorithm, t-distributed stochastic neighbour embedding (t-SNE) is applied to reduce the dimension of learned features for visualization. The features in the 5th convolutional layers are used for analysis, since this layer is the end of feature descriptor in traditional architecture and the learned features should be highly abstract and separable. Taking the data from 2009 PHM as an example, the results of feature visualization for CNN and MCNN versus small numbers of training samples (128, 256, 512) are shown in Fig. 12. It is clear that the features learned by MCNN are well clustered compared with the counterpart of CNN cases. For different number of samples, the features of CNN are mixed and overlapping, such as the spur 7(orange) and spur8 (red) and the traditional CNN fails the classification of these two kinds of faults. Again, the visualized results tell that the proposed MCNN features remarkable feature representation ability with limited number of training samples.

In addition, the training time for each training epoch and the memory usage during the training progress of the 2 datasets with different methods are recorded and presented in Table 3. As exhibited, for each training epoch, MCNN utilizes less computational time as well as low memory footprint. Compared to CNN, the computational times per training epoch of MCNN have been reduced by 0.168 seconds and 0.193 seconds. The memory footprint of MCNN have been reduced by 31.5MB and 28MB respectively in the two datasets.

5. Conclusion

In this work, a modified convolutional neural network that replacing the fully-connected layers with the global average pooling scheme, is proposed to reduce the number of trainable model parameters. The improved architecture possesses higher precision, less computational burden, superior generalization ability with limited training samples. Moreover, a MCNN-based intelligent fault diagnosis framework is presented. In order to assess the performance of the proposed method, the case studies on two industrial gearboxes are conducted from three aspects including the classification accuracy, the features visualization and the computational efficiency. These results demonstrate the

Table 3. Time cost and memory footprint by different approaches

Approaches	PHM2009 dataset		DDS dataset	
	Time(sec/epoch)	Memory (MB)	Time(sec/epoch)	Memory (MB)
CNN	0.915	383.1	0.873	232
CNN+l2	0.961	388.6	0.905	236
CNN+BN	0.914	410.3	0.909	243
MCNN	0.747	351.6	0.68	204

superiority of MCNN, compared to shallow machine learning methods i.e. SVM and RF and other popular deep learning approaches, such as CNN and its variant methods. Specifically, MCNN achieves the 51.3% and 24.6% improvements in the aspect of classification accuracy for two dataset with limited training data, i.e., 128 training samples. The impressive performances, achieved by the MCNN, show a broad prospect for intelligent fault diagnosis in the industrial gearbox.

Acknowledgement

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by the National Key Research and Development Program of China (Project No. 2017YFC0108401) and National Natural Science Foundation of China (5

References

1. Cai B P, Huang L, Xie M. Bayesian networks in fault diagnosis. IEEE Transactions on industrial informatics 2017; 13(5): 2227 - 2240, https://doi.org/10.1109/TII.2017.2695583.
2. Cai B P, Liu Y H, Fan Q ,Zhang Y W, Liu Z K, Yu S L, Ji R J. Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network. Applied Energy 2014; 114: 1-9, https://doi.org/10.1016/j.apenergy.2013.09.043.
3. Cai B P, Liu H L, Xie M. A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks. Mechanical Systems and Signal Processing 2016; 80: 31-44, https://doi.org/10.1016/j.ymssp.2016.04.019.
4. Cai B P, Shao X Y, Liu Z K, Kong X D. Remaining useful life estimation of structure systems under the influence of multiple causes: Subsea pipelines as a case study. IEEE Transactions on Industrial Electronics 2019; 99:1-1, https://doi.org/10.1109/TIE.2019.2931491.
5. Chen D, Yang S, Zhou F. Transfer learning based fault diagnosis with missing data due to multi-rate sampling. Sensors 2019; 19(8):1826, https://doi.org/10.3390/s19081826.
6. Chen Yuejian, Liang Xihui, Zuo Ming J. Sparse time series modeling of the baseline vibration from a gearbox under time-varying speed condition. Mechanical Systems and Signal Processing 2019; 134: 106342, https://doi.org/10.1016/j.ymssp.2019.106342.
7. Chen ZY, Gryllias K, Li WH. Mechanical fault diagnosis using Convolutional Neural Networks and Extreme Learning Machine. Mechanical Systems and Signal Processing 2019; 133: 106272, https://doi.org/10.1016/j.ymssp.2019.106272.
8. Duan R, Lin Y, Zeng Y. Fault diagnosis for complex systems based on reliability analysis and sensors data considering epistemic uncertainty. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 558-566, https://doi.org/10.17531/ein.2018.4.7.
9. Feng K, Wang K, Ni Q. A phase angle based diagnostic scheme to planetary gear faults diagnostics under non-stationary operational conditions. Journal of Sound and Vibration 2017; 408:190-209, https://doi.org/10.1016/j.jsv.2017.07.030.
10. Feng Z P, Zhu W P, Dong Zhang. Time-Frequency demodulation analysis via Vold-Kalman filter for wind turbine planetary gearbox fault diagnosis under nonstationary speeds. Mechanical Systems and Signal Processing 2019; 128: 93-109, https://doi.org/10.1016/j.ymssp.2019.03.036.
11. Han T, Jiang D, Qi Z, Lei W, Kai Y. Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. Transactions of the Institute of Measurement and Control 2018; 40: 2681-93, https://doi.org/10.1177/0142331217708242.
12. Han, T, Jiang, D, Sun, Y, Wang, N, Yang, Y. Intelligent fault diagnosis method for rotating machinery via dictionary learning and sparse representation-based classification. Measurement 2018; 118:181-193, https://doi.org/10.1016/j.measurement.2018.01.036.
13. Han T, Liu C, Wu Lj, Sarkar S, Jiang DX. An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems. Mechanical Systems and Signal Processing 2019; 117:170-187, https://doi.org/10.1016/j.ymssp.2018.07.048.
14. Han T, Liu C, Yang WG, Jiang DX. Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application. ISA Transactions 2019, https://doi.org/10.1016/j.isatra.2019.08.012.
15. Han T, Liu C, Yang WG, Jiang D. Learning transferable features in deep convolutional neural networks for diagnosing unseen machine conditions. ISA Transactions, https://doi.org/10.1016/j.isatra.2019.03.017.
16. Huang W Y, Cheng J H. An improved deep convolutional neural network with multi-scale information for bearing fault diagnosis. Neurocomputing 2019, https://doi.org/10.1016/j.neucom.2019.05.052.
17. Jiao J, Zhao M, Lin J. Deep Coupled Dense Convolutional Network with Complementary Data for Intelligent Fault Diagnosis. IEEE Transactions on Industrial Electronics 2019; 66(12): 9858 - 9867, https://doi.org/10.1109/TIE.2019.2902817.
18. Kaluer S, Fekete K, Jozsa L, Klai Z. Fault diagnosis and identification in the distribution network using the fuzzy expert system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 621-629, https://doi.org/10.17531/ein.2018.4.13.
19. Lei J, Liu C, Jiang D. Fault diagnosis of wind turbine based on Long Short-Term memory networks. Renewable Energy 2019; 133: 422-432, https://doi.org/10.1016/j.renene.2018.10.031.
20. Lin M, Chen Q, Yan S C, Network in Network, Neural and Evolutionary Computing. arXiv:1312.4400.
21. Liu H , Zhou J , Zheng Y. Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders. ISA Transactions 2018; 77: 167-178, https://doi.org/10.1016/j.isatra.2018.04.005.
22. Liu XL, Zhang XY, Wang LY. Fault Diagnosis Method of Wind Turbine Gearbox Based on Deep Belief Network and Vibration Signal. Society of Instrument and Control Engineers of Japan.
23. PHM, Phm data challenge 2009., https://www.phmsociety.org/competition/PHM/09, 2009.
24. Rezaei M, Yang H J, Meinel C. Deep Neural Network with l2-norm Unit for Brain Lesions Detection. arXiv:1708.05221.
25. Shao H D, Jiang H K, Zhao K. A novel tracking deep wavelet auto-encoder method for intelligent fault diagnosis of electric locomotive bearings. Mechanical Systems and Signal Processing 2018; 110: 193-209, https://doi.org/10.1016/j.ymssp.2018.03.011.
26. Shao SY, Wang P, Yan R Q. Generative adversarial networks for data augmentation in machine fault diagnosis. Computers in Industry 2019; 106: 85-93, https://doi.org/10.1016/j.compind.2019.01.001.
27. Sikora M, Szczyrba K, Wróbel, Michalak M. Monitoring and maintenance of a gantry based on a wireless system for measurement and analysis of the vibration level. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(2): 341-350, https://doi.org/10.17531/ein.2019.2.19.

28. Tang G J, Pang B. Gearbox Fault Diagnosis Based on Hierarchical Instantaneous Energy Density Dispersion Entropy and Dynamic Time Warping. Entropy 2019; 21(6): 593, https://doi.org/10.3390/e21060593.
29. Tyagi S, Panigrahi S K. A Hybrid Genetic Algorithm and Back-Propagation Classifier for Gearbox Fault Diagnosis. Applied Artificial Intelligence 2017; 1-20, https://doi.org/10.1080/08839514.2017.1315502.
30. Verstraete D, Ferrada A, Droguett E.L, Meruane V, Modarres M. Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings. Shock and Vibration 2017; 2017: 1-17, https://doi.org/10.1155/2017/5067651.
31. Wang J, Li S, An Z. Batch-normalized deep neural networks for achieving fast intelligent fault diagnosis of machines. Neurocomputing 2019; 329: 53-65, https://doi.org/10.1016/j.neucom.2018.10.049.
32. Wang K S, Heyns P S. Application of computed order tracking, Vold-Kalman filtering and EMD in rotating machine vibration. Mechanical Systems and Signal Processing 2011; 25(1): 416-430, https://doi.org/10.1016/j.ymssp.2010.09.003.
33. Wang K S, Heyns P S. The combined use of order tracking techniques for enhanced Fourier analysis of order components. Mechanical Systems and Signal Processing 2011; 25(3): 803-811, https://doi.org/10.1016/j.ymssp.2010.10.005.
34. Wen L, Gao L, Li X, Wen L, Gao L, Li X. A new deep transfer learning based on sparse auto-encoder for fault diagnosis. IEEE Transactions on systems, man, and cybernetics: systems 2017; 1-9.
35. Xu H and Chen G. An intelligent fault identification method of rolling bearings based on LSSVM optimized by improved PSO. Mechanical Systems and Signal Processing 2013; 35(1-2): 167-175, https://doi.org/10.1016/j.ymssp.2012.09.005.
36. Yang J, Guo Y Q, Zhao W L. Long short-term memory neural network based fault detection and isolation for electro-mechanical actuators. Neurocomputing 2019; 360: 85-96, https://doi.org/10.1016/j.neucom.2019.06.029.
37. Yu J, Xu YG, Liu K. Planetary gear fault diagnosis using stacked denoising autoencoder and gated recurrent unit neural network under noisy environment and time-varying rotational speed conditions. Measurement Science and Technology 2019; 30: 095003, https://doi.org/10.1088/1361-6501/ab1da0.
38. Zhang M, Wang K, Li Y. Motion Periods of Planet Gear Fault Meshing Behavior. Sensors 2018; 18(11), https://doi.org/10.3390/s18113802.
39. Zhang M, Wang K S, Wei D D. Amplitudes of characteristic frequencies for fault diagnosis of planetary gearbox. Journal of Sound and Vibration 2018; 432:119-132, https://doi.org/10.1016/j.jsv.2018.06.011.
40. Zhang W, Li X and Ding Q. Deep residual learning-based fault diagnosis method for rotating machinery. ISA Transactions 2018, https://doi.org/10.1016/j.isatra.2018.12.025.
41. Zhang W, Peng G, Li C, Chen Y, Zhang Z. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. Sensors 2017; 17(2): 425, https://doi.org/10.3390/s17020425.
42. Zhang X L, Wang B J, Chen X F. Intelligent fault diagnosis of roller bearings with multivariable ensemble-based incremental support vector machine. Knowledge-Based Systems 2015; 89: 56-85, https://doi.org/10.1016/j.knosys.2015.06.017.
43. Zhang Z Z, Li S M. General normalized sparse filtering: A novel unsupervised learning method for rotating machinery fault diagnosis. Mechanical Systems and Signal Processing 2019; 124: 596-612, https://doi.org/10.1016/j.ymssp.2019.02.006.
44. Zhao X L, Jia M P. A new Local-Global Deep Neural Network and its application in rotating machinery fault diagnosis. Neurocomputing, https://doi.org/10.1016/j.neucom.2019.08.010.

Yaxin LI
Kesheng WANG
Department of Mechanical and Electrical Engineering
University of Electronic Science and Technology of China
Chengdu, 610059, China

E-mail: yaxinli@std.uestc.edu.cn, keshengwang@uestc.edu.cn

Article citation info:
KONOWROCKI R, CHOJNACKI A. Analysis of rail vehicles' operational reliability in the aspect of safety against derailment based on various methods of determining the assessment criterion. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 73–85, http://dx.doi.org/10.17531/ein.2020.1.9.

Robert KONOWROCKI
Andrzej CHOJNACKI

ANALYSIS OF RAIL VEHICLES' OPERATIONAL RELIABILITY IN THE ASPECT OF SAFETY AGAINST DERAILMENT BASED ON VARIOUS METHODS OF DETERMINING THE ASSESSMENT CRITERION

ANALIZA NIEZAWODNOŚCI EKSPLOATACYJNEJ POJAZDÓW SZYNOWYCH W ASPEKCIE BEZPIECZEŃSTWA PRZED WYKOLEJENIEM W OPARCIU O RÓŻNE METODY WYZNACZANIA KRYTERIUM OCENY*

The article features the results of computer and experimental research on operational issues in the aspect of safety in relation to a freight wagon derailment on a railway track. It presents the knowledge regarding the methods of assessing the operational safety of rail vehicles on railroad tracks for the purpose of comparative analysis. The theoretical analyses were performed based on several methods that assess the safety of their derailments, qualifying for operational reliability, comparing them with the results obtained from experimental research. For the purpose of the research, a computer model of rail vehicle- railway track was created. It took into consideration dynamic parameters of elements used in the real track and rail vehicle. The results obtained from theoretical analyses were validated with experimental tests carried out on real objects (freight vehicle - test track, freight wagon - test rig). As part of the research, new test track geometry for testing rail vehicles was proposed. The results obtained in this way allowed estimating the conditions threatening the operation of a freight vehicle while running on the test rail infrastructure with different assessment criteria and to compare them.

***Keywords**: operational safety, rail vehicle dynamics, derailment, experimental tests, numerical investigations.*

W pracy pokazano rezultaty badań komputerowych i eksperymentalnych dotyczących zagadnień eksploatacji w aspekcie bezpieczeństwa w odniesieniu do wykolejenia wagonu towarowego na torze kolejowym. Przybliżono w nim stan wiedzy dotyczącej metod oceny bezpieczeństwa eksploatacji pojazdówa) szynowych na kolejowych liniach szynowych, w celu ich analizy porównawczej. W pracy wykonano analizy teoretyczne bazując na kilku metodach, które oceniają bezpieczeństwo ich wykolejenia, kwalifikujące się do niezawodność eksploatacyjnej, porównując je z wynikami otrzymanymi z badań eksperymentalnych. Na potrzeby przeprowadzanych badań powstał komputerowy model pojazd szynowy - tor kolejowy. Uwzględniał on parametry dynamiczne elementów zastosowanych w rzeczywistym torze oraz pojeździe szynowym. Otrzymane z teoretycznych analiz wyniki zwalidowano testami eksperymentalnymi wykonanymi na rzeczywistych obiektach (pojazd towarowy - tor testowy, wagon towarowy - stanowisko badawcze). W ramach badań zaproponowano nową geometrię toru testowego do badań pojazdów szynowych. Uzyskane wyniki pozwoliły określić stan zagrożenia eksploatacji wagonu towarowego podczas jazdy po testowej infrastrukturze szynowej przy różnych kryteriach oceny oraz je porównać.

***Słowa kluczowe**: bezpieczeństwo eksploatacji, dynamika pojazdów szynowych, wykolejenie, badania numeryczne i eksperymentalne.*

1. Introduction

The problems of safety and reliability of rail vehicles movement is constantly being developed in scientific research [41]. Processes aiming at increasing the level of operational reliability and safety are taken into consideration as early as their design. World trends in this issue require using theories of safety reliability and such vehicles' operational safety [29, 34, 39]. The theory of safety evolved in the 90s so as to counteract risks of failures or accidents which could lead not only to the disruption of a particular technical system functioning but the loss of health, human life or other damage as well [21]. As far as rail vehicles are concerned derailment is the most common kind of accident which causes at the same time the risk of infrastructure degradation, rolling stock or transported cargo damage, disruption of services and environmental damage (the transportation of hazardous materials) [16, 18]. Therefore the risk determination of such an event occurrence is vital during rail vehicles' tests.

Theoretical analyses and tests are carried out while designing of such vehicles in order to forecast the impact of vehicle parameters on a test track [1, 9, 28] as well as running safety and reliability connected with derailment during movement and monitoring [6, 7]. Then they are continued experimentally in acceptance tests during qualifications to place them in service or after extensive maintenance/modernization of rail vehicles in operation [11, 35]. Safety against derailment is one of primary criteria for assessing the reliability of rail vehicle operation.

Many researchers still deal with the topic of rail vehicle safety [10, 48, 49]. In many cases, the main mechanism causing train derailment is the loss of lateral stability of a railway vehicle [6, 24, 43]. This is caused by a rise in the lateral force value in the wheel-rail contact

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

zone. This could the consequence of various conditions resulting in a loss of lateral guidance provided by the track during normal vehicle operation. The following should be mentioned: climb of the wheel flange, extension of the rail gauge, rail inclination, track condition [49, 50] and reduction of lateral stiffness of the fastening system to sleepers [20].

Apart from theoretical research there are also conducted experimental freight wagons derailment tests [42, 50], where permissible relative wheelset unloading $\lim\Delta q$ was analyzed. The tests showed that unloading may be in the range of $0.62 \leq \lim\Delta q \leq 0.84$. In most cases safety analysis rely on assessment criteria which base on values of *Y/Q* derailment quotient, wheels unloading and their lift Δz. The review of the literature allows dividing the methodology of safety against railway vehicle derailment at different assessment criteria. The following should be mentioned here:

1. Nadal single-wheel *Y/Q limit criterion*, – binding for small running speeds on track curves [10].
2. *Weinstock axle-sum Y/Q limit criterion*, [48].
3. CHXI 50-millisecond time *limit*- Association of American Railroads (AAR) [12]
4. Y/Q time duration criterion - Japanese National Railways (JNR) [32].
5. Y/Q time duration criterion - Electro-Motive Division (EMD) of General Motors, Inc. [25].
6. Wheel climb distance criterion proposed by Transportation Technology Centre, Inc. (TTCI) [47].

There are several reasons for the risk of rail vehicle derailment to occur. One of the main derailment scenarios takes place when during a vehicle motion a high lateral force acting on the wheelset causes the wheel flange to come into contact with the rail. As a result of this contact, the wheel quickly climbs over the rail and when the maximum value of the flange angle is reached, the wheelset derails. The wheel climb causing derailment is connected with exceeding the limit value of the ratio of lateral force components *Y* to the vertical force *Q* at the wheel-rail contact, see Fig. 1. In such a case, the *Y/Q* force ratio is usually called the derailment quotient.

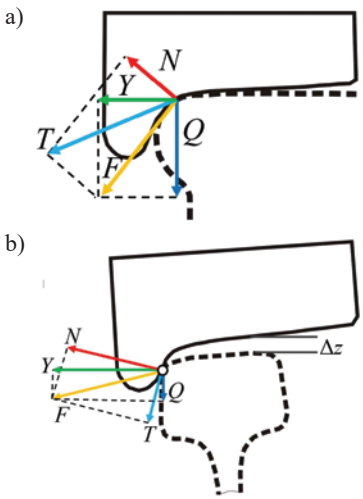


Fig. 1. Force components in the wheel–rail contact on the straight line (a) and on the curve (b): lateral forces (Y), vertical force (Q), normal force (N), lateral rolling friction force (F), flange tilt (γ), wheel lift (Δz)

The values of force components presented in Fig. 1 could be defined as follows:

$$\begin{aligned} Y &= F\cos(\gamma) + N\sin(\gamma) = Y_{cos} + Y_{sin} \\ Q &= -F\sin(\gamma) + Q\text{scos}(\gamma) = Q_{cos} + Q_{sin} \end{aligned} . \tag{1}$$

Nadal criterion [10] defining the derailment quotient is based on the value of friction coefficient μ between wheel and rail as well as flange tilt angle γ. Mathematically, they are determined by (2). This criterion is easy to implement and is therefore widely used to assess safety against derailment.

$$\frac{|Y|}{|Q|} < \frac{tg\gamma - \mu}{1 + \mu tg\gamma} . \tag{2}$$

In particular, the above criterion is applied to assess the risk of flange climb derailment in UIC 518 Code [46] and European Standard EN14363 [11]. The main modification adopted in these two documents is the requirement that the *Y/Q* quotient does not exceed the assumed critical value of 1.2 at a distance of 2 m in the vehicle's movement distance interval in the quasi-static tests.

Another standardized [17] index applied to assess rolling stock dynamics related to running safety on 1520 mm gauge tracks is also known. It is called the wheelset stability safety coefficient against derailment in the case of rolling of a wheel flange at/on the rail head and marked with k_z. It is determined from the formula (3). This coefficient is a mathematical modification of the formula (2) and its maximum permissible value for freight wagons is 1.3 [17].

$$k_z < \frac{tg\beta - f_{FR}\mu}{1 + f_{FR} \cdot tg\beta} \cdot \frac{P_v}{Y} \geq [k_z] , \tag{3}$$

where: β is the tilt angle of the conical part of the rim flange surface to the horizontal reference line, f_{FR} describes the slip friction coefficient in the wheel/ rail contact area, P_v and Y are respectively the vertical and horizontal force components of the wheel interaction on the rail.

The aim of this article is to analyze the safety assessment and thus the operational reliability of rail vehicles against derailment, as well as to propose a new test track geometry that takes into account track twist based on the bogie and vehicle base. Such a track allows determining safety indicators against derailment during one ride of a rail vehicle, without the need for additional tests at stationary stands described in the further section of this paper.

2. Methods of risk assessment and safety against derailment applied during experimental tests

New designs of rail vehicles that will be placed in service in the European Union should meet the essential requirements set out in the Technical Specifications for Interoperability (TSIs). One of the requirements is to check whether the vehicle can be safely operated on the tracks. Both the TSI on freight wagons and the TSI on locomotives and passenger vehicles consider the fulfillment of the requirements given in the EN 14363:2005 standard as proof of safety (the current edition of the standard is EN 14363:2016 [11] – adopted in Poland and marked as PN-EN 14363 + A1:2019-02). Before the adoption of the aforementioned standard, risk and safety tests against goods wagons derailment were carried out based on the requirements provided in the ERRI report (ORE) [40]. No separate requirements were developed to assess the running safety on the twisted track of other rail vehicles. Therefore, the report [40] was also used to determine the safety factor of driving on a twisted track of these vehicles. The standard [11] applies to both quasi-static tests (the speed of the tested rail vehicle does not exceed 10 km/h) and dynamic tests (provided for vehicles with a permissible speed above 60 km/h). This document lists 3 methods for testing operational safety on a twisted track.

Method 1 – vehicle investigations when running on a twisted track. In this case, a track with a constant curve radius *R*=150 m is used as the measuring test rig. Track twist is implemented by chang-

ing the height of the outer rail (positive and negative track cant). The track twist necessary for the test is 3 ‰ over a 30 m section. Moreover, the track construction should reflect the normal conditions of a typical track, taking into account the rail profile, track gauge and maintenance status. During the tests, there must be no lateral forces in the train composition and the vehicle itself must not be braked. The running on the curve is carried out by pushing or pulling the vehicle. The tests are conducted on dry rails so that the wheel/rail coefficient of friction is the highest. The previous document – ORE Report [40] regarding safety tests and dynamics of freight wagons running on a twisted track recommended so that the track be washed with a technical solvent before launching the tests. In the next stage of track preparation, it was required to sprinkle the rail with fine sand and then sweep it off the surface of the rail head. The track prepared in this way ensured a high wheel/rail coefficient of friction.

The European standard [11] lists the mathematical relations from which the twist required on the basis of the bogie and the vehicle base must be determined. If it exceeds 3 ‰, the vehicle must be prepared in a specific way. This can be done, for example, using shims placed under the vehicle suspension. The guidelines for calculating the thickness of the shims and their distribution are given in the annex of EN 14363 standard. Before starting the safety tests, the vertical wheel loads must be determined on the vehicle. Then the test should be planned so that the wheel with the least vertical force is a leading wheel, i.e. a wheel that attacks the outer rail on the measuring curve. If this guideline is not met, this should be done by placing additional shims under the suspension in a different location.

During the investigations, a minimum of 3 tests are required to run the vehicle on the curve at a constant speed not exceeding 10 km/h. The measured parameters include: guiding forces on the inner and outer wheel of a tested vehicle Y_i, Y_a, vertical wheel forces on the inner and outer wheels of the tested vehicle Q_i, Q_a, the angle of attack of the leading wheelset α, leading wheel climb Δz on the entire curve. The abovementioned parameters can be measured by devices placed in the track or on the vehicle. If the measuring devices are placed in the track, their location should be on a twisted track section. Detailed guidelines can be found in the EN 14363:2016 standard [11], and an example of the location of measuring points on the curve of the test track is demonstrated in Figure 6.

The leading wheel lift Δz must be registered in a continuous way. The measurements of forces (Y_i, Q_i) on the inner rail as well as the angle of leading wheelset α serve only to verify the wheel/rail coefficient of friction. On the basis of measured values of vertical Q_a and lateral Y_a forces in selected measuring points on the track curve, running safety quotients $(Y/Q)_a$ were determined. The maximum value of the running safety ratio on a twisted track $(Y/Q)_{a,max}$ was assessed. A vehicle is considered safe if condition (4) is met:

$$\left(\frac{Y}{Q}\right)_{a,lim} \leq \left(\frac{Y}{Q}\right)_{lim} . \tag{4}$$

According to Nadal criterion, for the wheel of flank cant of 70° and coefficient of wheel/ rail friction $\mu = 0.36$ this condition is $(Y/Q)_{lim} = 1.2$. If the limit value $(Y/Q)_{lim}$ is exceeded, the wheel climb over the rail head should be checked. If the relation $\Delta z_{max} \leq \Delta z_{lim}$ (where: $\Delta z_{lim} = 0.005$ m) is met, it means that the vehicle has actually not derailed. Therefore this vehicle can be considered safe if it fulfills the following conditions: the flange angle does not exceed 70° in any profile location, then it must be documented that outer rail is dry and it is free of grease or foreign matter, the test has been conducted at least 3 times and in each case the condition $\Delta z_{max} \leq \Delta z_{lim}$ has been met.

Therefore, the final criterion for assessing safety of a vehicle running on the track is the fulfillment of the leading wheel lift criterion $\Delta z_{max} \leq \Delta z_{lim}$.

Method 2 of the rail vehicle derailment risk analysis relates to tests carried out on a test rig simulating the impact of a twisted track on a vehicle and the passage of a test vehicle over a test track without twist. In order to assess the running safety on a twisted track, based on Method 2 - described in the standard [11], two test rigs should be used. The first are special twisting stations, where the minimum wheel force Q_a is determined during running on the twisted track. The second test rig is a track with a radius of $R = 150$ m without cant. At this track, the wheelset maximum guiding force Y_a is determined. In order to determine limit twist, the same relations as described in Method 1are applied. Due to the fact that the wheel lift up to $\Delta z = 0.005$ m is actually permissible, the dependencies on the limit twists are reduced on the test rig.

The test rig used for measuring wheel force Q_{jk} (j- means wheel set number, k- vehicle side) must be equipped with devices for lifting and lowering them. Independent wheel displacement should be carried out on at least two wheelsets of one bogie. While twisting, Δz_{jk} wheel displacement and the wheel force Q_{jk} of all the wheels are measured continuously. Based on the data processing, taking into account the forces caused by the combined twist of the body and the bogie, the eccentricity of their centers of gravity, including friction and deviations, the minimum vertical wheel force $Q_{jk,min}$ is determined. Figure 2 shows an example of wheel displacement during vehicle twist, performed to determine the value of individual wheel forces on rail tracks. This process carried out on the test rig simulates the change in wheel forces during the vehicle running on a twisted track.

Fig. 2. Positions of suspension beams of measurement modules while conducting tests of freight wagon wheel forces – freight wagon twisting, (a) lifting the wheel over the zero level of the rail head (b) lowering the wheel in relation to the zero level of the rail head

When the maximum guiding force Y_a is determined, the track should consist of a straight section and a curve of a radius R = 150 m. The test rig should not have a transition curve, cant or turn. As in Method 1, the track construction should reflect typical track normal conditions taking into account the rail profile, track gauge and maintenance status. Test drives should be planned so as the wheel of minimal vertical force be the leading wheel. In this method, the longitudal force generated on the train composition must be levelled and the tested vehicle must not be braked. The tests must be carried out at least 3 times at a speed not exceeding 10 km/h. The values measured in this test on a curve include: leading forces on inner and outer wheel of the tested vehicle Y_i, Y_a, the vertical wheel force of on the inner wheel of the tested vehicle Q_i and the angle of attack of a leading wheelset α.

The above parameters can be measured by devices placed in the track or on the vehicle. If the measuring devices are placed in the track, their location is in two zones. The first zone is at the beginning of the curve over a distance of more than 3 m up to $2a^*$ ($2a^*$ - bogie centre distance or distance of the axles in non-bogied vehicles). The location of this zone provides the forces measurement during the rotation of the bogie against the body, which is essential for such vehicle constructions. A minimum of 3 measuring points are provided in this zone. The next measuring zone should be located so that the entire vehicle is in the curve. The beginning of the zone should be placed at a distance above $2a^+ + 2a^*$ (with $2a^+$ - as the bogie wheel base) from the beginning of the curve. This zone should also contain a minimum of 3 measurement points.

During the application of this test method, the guiding forces Y_i and Y_a should be recorded for each measurement position. Their assessment should be performed with the use of mean value $Y_{i,med}$ and $Y_{a,med}$ from measurement points separately for each measurement zones. The force direction Y_i in most cases is opposite to force Y_a. Bearing in mind the measured parameters such as (Y_i, Q_i) forces on an inner rail and leading wheelset angle of attack α it is assumed that wheel/rail friction is close to the friction between wheel and rail limit value. Therefore the track should be prepared in a similar way as described in Method 1.

The assessment of meeting the safety requirements and thus minimizing the risk of derailment, while running on a twisted track should be carried out for each wheelset. In this case, in compliance with EN 14363:2016 standard [11], formulas (5) and (6) defining the value of derailment ratio and its limit are applied:

$$\left(\frac{Y}{Q}\right)_{ja} = \frac{Y_{ja,med}}{Q_{jk,\min} + \Delta Q_{jH}}, \tag{5}$$

where: $Y_{ja,med}$ is the quasi-static guiding force determined on the basis of the vehicle running on a curve with a radius of R =150 m, $Q_{jk,min}$ denotes the smallest vertical wheel force calculated on the basis of a twist test, ΔQ_{jH} represents the change in the vertical wheel force due to the moment of the guiding forces. The criterion for assessing the derailment risk in this case takes the form of a formula:

$$\left(\frac{Y}{Q}\right)_{ja} \le \left(\frac{Y}{Q}\right)_{lim}, \tag{6}$$

according to Nadal equation, for a flange angle of 70° and the flange coefficient of friction μ = 0.36. The limit value of derailment quotient is $(Y/Q)_{lim}$ = 1.2.

Method 3, covered by standard [11] refers to tests of vehicles on twist test rig and yaw torque test rig. It can be applied for conventional technology vehicles, i.e. those which operate in normal conditions and they or their construction components connected with their running properties comply with the recognized state of knowledge. Thus they include two-axle bogies, two bogies per vehicle and flange angles of wheels γ between 68° and 70°. Therefore this method cannot be applied for combined-bogie vehicles and rail vehicles fitted with three-axle bogies.

In order to assess the operational safety on a twisted track and the risk of derailment, based on Method 3, two test rigs should be used, i.e. a test rig for measuring wheel force and a test rig for measuring the bogie rotational resistance torque relative to body. The rig used to measure wheel force, as in Method 2, should be equipped with devices for lifting and lowering the wheels. Independent wheel displacement must be carried out on at least two wheelsets of one bogie. The track twist required in the tests is determined from the same formulas as in Method 1. Therefore, the range of wheel displacement during wheel force measurements in Method 3 is greater than the range determined in Method 2. The vehicle wheel force algorithm is the same as in Method 2. During twisting displacement of wheels Δz_{jk} and the wheel force Q_{jk} of all wheels are measured continuously. Based on these data, the vertical wheel force of the tested wheel Q_0 on a levelled track and the wheel force drop ΔQ caused by maximum twist are determined.

The second stage of tests related to measuring the body-to-bogie yaw torque is carried out on a stationary test rig. This position allows the bogie to rotate left and right relative to the body by a given angle. The required rotation speed of the rig is constant and amounts to 1°/s within 75% of the bogie rotational angle relative to the body. Due to the load on the wagon during operation, the body-to-bogie yaw torque should be measured for the empty and loaded vehicle. During the test, the bogie should be connected to the vehicle body by all anticipated connections. The essence of this method is to determine the ratio of the leading wheel unloading to the vertical wheel force in the absence of twist and the X factor characterizing the behaviour of the bogie on small radius curves. Permissible values of the above parameters are described by formulas (7) and (8). In compliance with EN 14363: 2016 [11], a vehicle is considered safe if it meets two criteria at the same time:

$$\frac{\Delta Q}{Q_0} \le 0.6, \tag{7}$$

where: ΔQ is the deviation from Q_0 at maximum twist conditions, Q_0 denotes the wheel force for the tested wheelset on levelled track and the bogie rotational resistance factor X for freight wagons depends on axle load. Factor X should be determined depending on:

$$X = \frac{M_{z,Rmin}}{2a^+ 2Q_0}, \tag{8}$$

where: $M_{z,R\ min}$ is the value of the bogie torque relative to the body evaluated at $\psi = a^*/R_{min}$, 2a+ denotes distance of wheelsets in a bogie (bogie wheelbase is 1.8 m), $2Q_0$ is wheel force of a tested wheelset. In case of freight wagons, the factor X value is determined from the diagram presenting factor X limit value depending on axle force on a track $2Q_0$ (Fig. 5).

3. Experimental investigation

3.1. Tested object

A typical Eanoss series coal wagon equipped with two standard Y25 bogies was used for experimental research. The tested vehicle is intended for transporting aggregate, coal and bulk materials. It can operate on 1.435 m gauge tracks. The geometrical parameters of the

wagon are: total length with buffers LUP = 14.04 m, maximum width 3.038 m, while the maximum height is 3.43 m. The weight of the tested wagon was 20.3 t. The wagon was designed for a maximum wheelset load of 22.5 t, i.e. up to a gross weight of 90 t. If empty, the wagon can be operated at a maximum speed of 120 km/h, while if loaded - up to 100 km/h.

The parameters necessary to prove that the wagon can be safely operated are: distance between bogie centres $2a^{*}$ = 9.0 m, wheel base distance in a bogie $2a^{+}$ = 1.8 m.

3.2. Description of the test metod

In order to check if the tested Eanoss series coal wagon can be safely operated, Method 3 specified in EN 14363:2016 standard was used. Two stationary test rigs were used during the experimental tests on the freight wagon. In the first stage of testing it was a test rig for measuring wheel force, while in the second stage for measuring a test rig for measuring the bogie-to- the body yaw torque. Wheel forces measurements were carried out at the TENSAN-PLW test rig, which is capable of independent displacement of individual wheels. At the test rig in question, the Railway Research Institute has its own software, which is used to perform vertical forces of individual wheels during the tests, during which their forces and vertical displacement are recorded. The software used for testing was prepared on the basis of vehicle twist guidelines provided in EN 14363:2016 standard.

The TENSAN-PLW test rig is fitted with specialized measuring modules, which include swingarm pivot mounted in the track axis as seen in Figure 2. Their displacement, relative to the zero track level, is forced by hydraulic cylinders. The location of the measuring modules and their length enables placing each vehicle wheel on them separately, and thus allows individual measurement of their force (Fig. 2). The total length of the test rig is 22.22 m. The range of maximum vertical force on a single swingarm is Q = 200 kN and the swingarm displacement is within the range Δh = ± 0.07 m. In the case of wheel force measurements of articulated vehicles whose length is greater than the test rig length, the track in front and behind of the test rig is levelled, which allows obtaining a vertical difference in the rails alignment.

3.3. Experimental results

During the tests of the Eanoss series coal wagon, wheel force was measured while lifting and lowering individual wheels from a neutral track level. Two twist tests were carried out, i.e. a test on a fixed wheel set (the tested wheel set which is subject to assessment is not displaced during tests, the wheels in a second bogie are displaced) and a test on a movable wheel set (the tested wheel set is displaced). In the article, in order to limit the number of pages, only the results of the analysis of the wheel forces of the first right front bogie are presented - diagrams in Figure 3. Based on them, the minimum and maximum wheel force ($Q_{0,jkx\ min}$ and $Q_{0,jkx\ max}$), occurring on a levelled track were determined while twist tests. From the arithmetic mean obtained from these forces, the nominal wheel force was determined without the effect of friction hysteresis, which in the analyzed case was $Q_{0,jkx}$=24.12 kN. These tests also allowed determining the torsional stiffness of the tested vehicle measured on the base of distance between the spherical pivot $C^{*}_{tA_{jk}}$ and the axle in the bogie $C^{+}_{tA_{jk}}$. The above parameters were used to determine the minimum wheel force during the vehicle's ride on the track with a given twist. At a further stage of tests, they were used in theoretical research by implementing them in the numerical model of a freight vehicle presented in Chapter 4 of this article.

The test rig for body-to-bogie yaw torque, belonging to the Railway Institute, was used for testing the aforementioned wagon in the next stage of tests. During tests, the maximum rotation angle of the rig was ψ= ± 10°, whereas the rotation speed was set to three values, i.e. 0.2, 0.6, 1.0 [1°/s]. Finally, in compliance with the standard

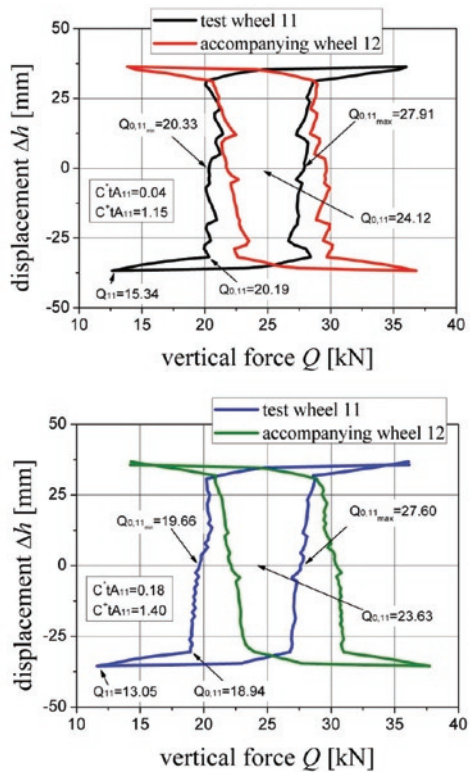


Fig.3. Measurements of the wheel force on the test rig, measurement on fixed (top) and moving (bottom) wheelset

[11], the results recorded at the rotation speed equal to 1.0 [1°/s] were taken into account for the assessment. Other measurements, which are not subject to assessment, are only used to obtain the appropriate interaction of friction elements to connect the proper measurement of the bogie with the wagon's body. The test was carried out on an empty and loaded wagon, i.e. with a gross load of 89.7 t. At each speed rotation of the bogie relative to the body and for each load of the wagon, one measurement of the body-to-bogie yaw torque was carried out. The results of the resistance torque on the spherical pivot obtained during experimental and simulation tests, with two variants of the wagon load and a rotation speed of 1.0 [1°/s] are shown in the diagrams (Fig. 4).

Measurement of bogie-to-bogie yaw torque relative to freight wagon body, freight wagon empty (left), freight wagon loaded (right) measurement velocity $V_{obr.}$ = 1.0 [°/s]

The numerical values of the torque at several rotational speeds and different vehicle loads are presented in Table 1. These results were obtained using the assessment criteria according to the limit value of the X factor (Fig. 5). In both load cases, the torque increases as the speed increases during rotation. This indicates a significant impact of degressive friction between the bogie and the wagon body. The numerical model of the freight wagon used for the simulation tests shown in Figures 4 is described in the further part of the paper.

The aforementioned test rig for dynamic testing of force values in the wheel-rail contact zone was a test track arranged on a radius R = 150 m curve over a length of 95 m S-49 rails with a 1:40 inclination, spring-mounted to INBK-7 concrete sleepers with the SB-3 fastening system were used in this track. The spacing of the sleepers is 0.60 m, and the curve uses a nominal track extension of e = 0.005 m. Force measurement is performed at the rail level using DC double strain gauges (Fig. 6a) stuck on both rails, located in measuring cross sections 1, 2, 3, 4, 5 and 6 (Fig. 6b). The distance between the measuring cross-sections is determined based on the EN 14363 standard [11]. The value of the wheel force Q is determined from the linear relation

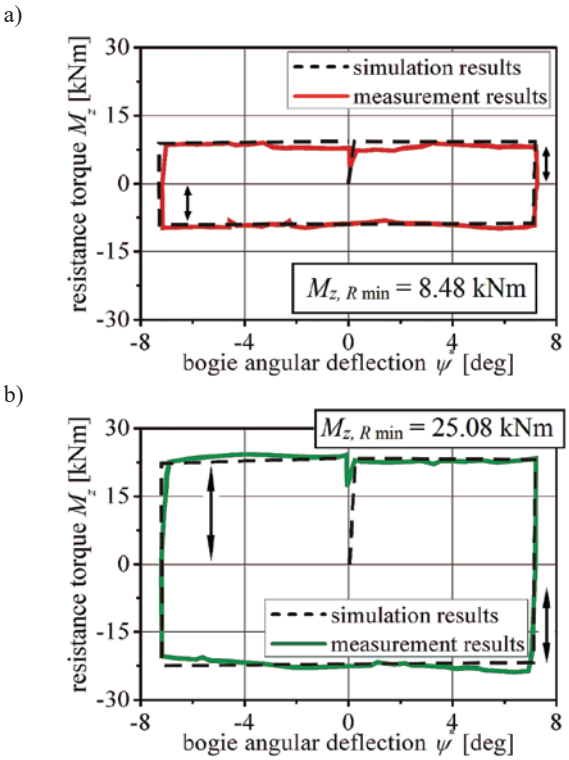


Fig. 4. Limit value of X factor for freight vehicles depending on axle force on the track

Table 1. Measurement results of a freight wagon torque

Bogie type	Wagon loading state	Value of yaw torque $M_{z,\,R\,\min}$ [kNm] at different measurement velocities		
		$V_{obr.}$ = 0.2 [°/s]	$V_{obr.}$ = 0.6 [°/s]	$V_{obr.}$ = 1.0 [°/s]
Y25	empty	9.54	8.72	8.48
	loaded	26.07	25.94	25.08

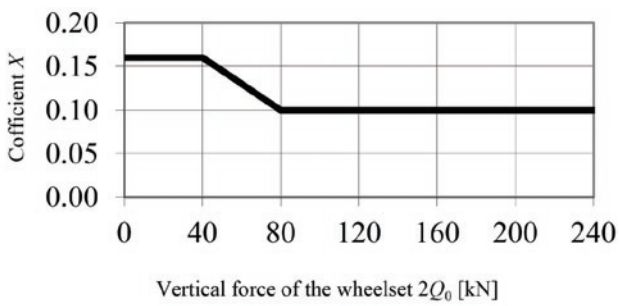


Fig. 5. Location of strain gauges on a rail in a single measuring point (a) and positions of all measurement cross sections 1, 2, 3, 4, 5 and 6 on a tested track curve

of the signals values from the measuring system (strain gauges R_1 and R_2 on the rail neck), which are strengthened by measuring amplifiers dedicated to them.

Linear relation of the measured signal corresponding to the horizontal wheel force Y, correlated with the impact of the vertical wheel force Q impacting the measuring point was determined on a calibration stand and recorded with the formula:

$$\begin{cases} Y_0 = Y - K_0 Q \\ Y_u = Y - K_u Q \end{cases}, \tag{9}$$

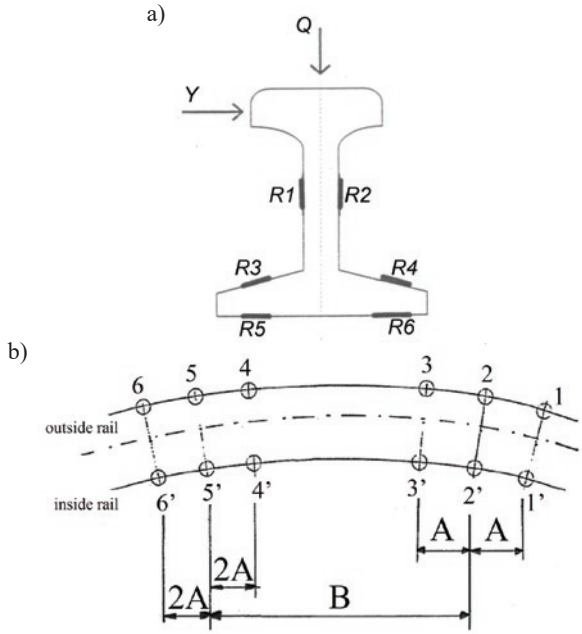


Fig. 6. Lateral forces Y in two measurement cross-sections (2 and 3) when rolling the freight wagon on the test track

where: Y_0 means is the measurement from a double strain gauge stuck to the upper part of the rail foot (strain gauges R_3 and R_4), Y_u denotes the strain gauge from the bottom of the foot (strain gauges R_5 and R_6), Q describes the vertical wheel force interaction on the rail, whereas K_0 and K_u are amplification factors of measuring amplifiers. From the above formulas we obtain (10) determining the value of lateral force Y at the measuring point:

$$Y = \frac{Y_0}{\left(1 - \frac{K_0}{K_u}\right)} - \frac{Y_u}{\left(\frac{K_u}{K_0} - 1\right)}. \tag{10}$$

Experimental testing of the Y/Q quotient was carried out when moving a freight wagon along the track at a speed of v=5 km/h. Time courses of the Y and Q forces values measured in cross-sections 1, 2, 4, 5 on the outer rail of the track curve and determined from 9 and 10 are presented in diagrams in Figures 7 ÷ 9. Based on the data obtained from each measuring point, the maximum values of derailment quotient for each wheel of a rail vehicle were determined, which are given in Table 2. In this table, the individual wheels of the tested vehicle are marked with the symbol W with indexes i, j. Index i denotes the vehicle axle number, whereas index j the wheelset side (j=1 right, j=2 left).

Based on the analysis of the results recorded in the given measuring cross-sections during the test runs, the wheelsets of the vehicle bogies were observed to transitionally unload the track laterally. This was caused by different angles of attack between the first and second wheelsets of the given bogie. In the case under analysis, when entering the curve, differences in lateral forces of single bogie sets were notices at the level of 50% and 36%, respectively in the first and second bogies of the tested wagon (Fig. 7). When exiting the curve, the differences in lateral forces between the sets of a single bogie are much larger and amount to over 60% (Fig. 8). This behaviour of wheel sets also resulted in a reduction in the wheel force of the second wheel sets on the track, which in the analyzed case ranged from 7% to 15% compared to the first set force (Fig. 9).

This behaviour significantly increases the value of the derailment ratio. A similar observation was obtained from the analysis of numeri-

Table 2. Derailment ratio values determined from experimental tests on the track

Wheel W_{ij}	11	12	21	22	31	32	41	42
Y/Q [kN]	0.73	0.72	0.21	0.24	0.59	0.69	0.66	0.62

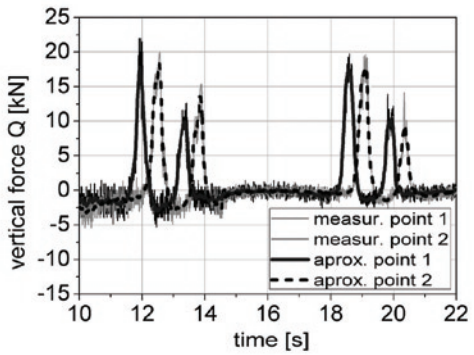


Fig. 7. Lateral forces Y in two measurement cross-sections (4 and 5) when rolling the freight wagon on the test track

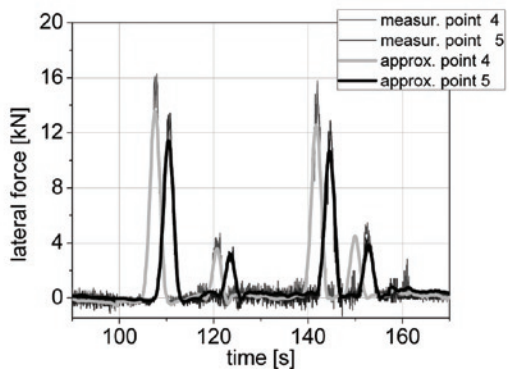


Fig. 8. Vertical forces Q in two measurement cross-sections (4 and 5) when rolling the freight wagon on the test track

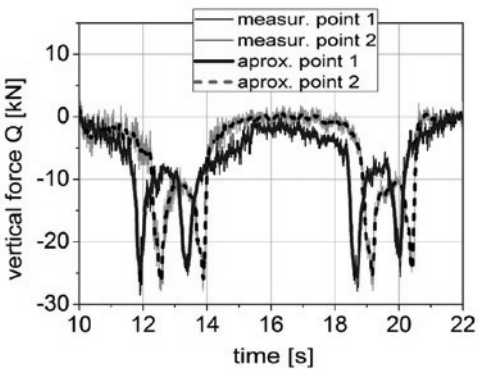


Fig. 9. Vertical forces Q in two measurement cross-sections (4 and 5) when rolling the freight wagon on the test track

cal results performed on the model of the analyzed vehicle described in the next section of the article. The phenomena of the dominant lateral interaction of the attacking wheelset (the first wheelset in the bogie) while negotiating the curve often occurs in rail vehicles [33, 43]. In such conditions, bogies hunting usually occur, and then at the highest lateral instability amplitude the longitudinal symmetry axis of the bogie is deviated from the longitudinal symmetry axis of the track by a certain angle α, called the angle of attack [3, 6, 30]. In this case, the wear of wheelsets and the running surface in the form of polygonization and corrugation also increases [2].

4. Numerical analysis of phenomena accompanying the risk of derailment

4.1. Numerical model of the rail vehicle

In order to determine theoretically the values of forces in the zones of contact between the wheels of the vehicle and the track rails, a model of the wagon mounted on Y25 series bogies was created. The physical model of the vehicle in question was treated as a system of rigid bodies joined together by means of elastic-damping elements (Fig. 10a and 10b). This approach in modeling is called the multi body system method [13,44] and is very often used by researchers in their own codes to analyze rail vehicle dynamics [3,26,45] and in commercial programs, i.e. Vampire, ViGrade/VI-Rail, Autodyn, Simpack, UM Loco. In this method, vehicle structural elements are treated as non-deformable bodies, while suspension elements describe flexible elements [4,8]. Limits of these bodies' motion result from the imposed integrable holonomic constraints [37].

The wagon model considered in the paper was divided into three basic elements, which are wheelsets, bogie frames and a vehicle body. The body was divided into two separate blocks in order to include the experimentally determined body torsional stiffness $K\theta_r$ in the vehicle model (Fig. 4). The wheelset was described as a block with three degrees of freedom, where lateral displacements were marked with the symbol (y_i), the angle of attack (ϕ_i) and the galloping (ϕ_i). The index i used in the description of the symbols corresponds to individual wheelsets, assuming the values of i=1, 2, 3, 4. The bogie frame is represented by a block with five degrees of freedom, which corresponds to lateral displacement (y_{rj}), vertical displacement (z_{rj}), yaw (ϕ_{rj}), angle of sway (θ_{rj}) and galloping angle (ϕ_{rj}). In this case, the index is j = 1.2, as there are two bogies. In the case of the vehicle body in question, the blocks describing them have five degrees of freedom. These include lateral displacement (y_{nk}), vertical displacement (z_{nk}), angle of yaw (ϕ_{nk}), angle of sway (θ_{nk}), galloping angle (ϕ_{nk}), at k=1.2 as the body is divided into two blocks. Summing up, the numerical model of the rail vehicle created for research purposes had 64 degrees of freedom.

In the case under consideration, the mathematical model of the freight wagon was described by a system of differential equations, which were derived using Lagrange equations of the second kind. In this approach, generalized coordinates take the form of linear displacements or angles of rotation. The motion of such a vehicle is described by ordinary differential equations of the second order. Assuming that the oscillations of individual blocks of the model relative to the reference system are small, such a system can be written as a linearized system of equations (11) written in the matrix form [14]:

$$[\mathbf{M}\, \mathrm{d}^2/\mathrm{d}t^2 + \mathbf{C}\, \mathrm{d}/\mathrm{d}t + \mathbf{K}]\cdot\mathbf{q}=\mathbf{F} \tag{11}$$

where: $\mathbf{q}=\{y_i, y_{rj}, y_n, z_{rj}, z_n, \varphi_i, \phi_{rj}, \phi_n, \phi_i, \theta_{rj}, \theta_n\}^T$ is vector of generalized coordinates, **M** denotes symmetrical inertia matrix,**C** is damping matrix, **K** is stiffness matrix, **F** denotes vector of forces and d/dt is differential operator.

Applying Newton-Raphson methods obtained the system of linear of linear algebraic equations, which is solved in each iteration with a time step Δt=0.001 s.

Detailed inertia parameters of the model's elements are presented in Table 3. The wheelbase of the wheelset was 1.5 m, the wheel radius R=0.42 m, the bogies distance of 9 m and the wheelset distance of 1.8 m, the distance of the slide side bearings is equal 1.7 m. The primary suspension stiffness $k_{Ix}=k_{Iy}$=4000 kNm and k_{Iz}=3950 kNm respectively, the torsion pivot stiffness was assumed at the level $k_x=k_y=k_z$ =10 MNm, the lateral and vertical stiffness of the slide side bearings is k_{sl_z}=350 kNm and k_{sl_z}=500 kNm, respectively. The parameters

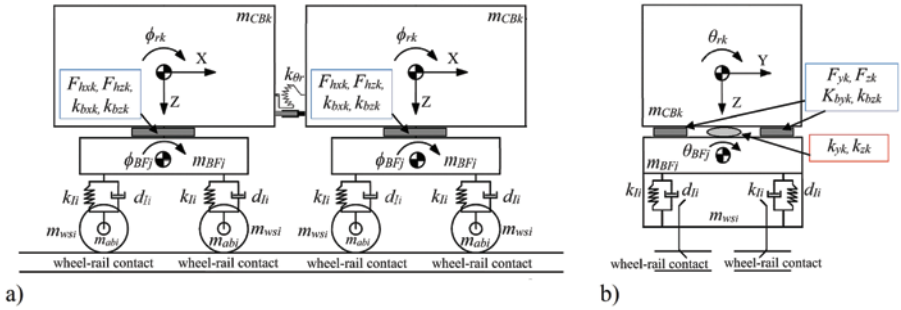


Fig. 10. Physical model of the analyzed vehicle (a) side view, (b) front view of the vehicle

Table 3. Inertia parameters of the freight wagon model

Body	Mass [kg]	I_{xx} [kg×m^2]	I_{yy} [kg×m^2]	I_{zz} [kg×m^2]
body m_{CB}	16000.00	47500.00	51000.00	50050.00
bogie m_{BF}	2000.00	1975.00	2850.00	1560.00
axle box m_{ab}	25.00	10.00	10.00	10.00
wheelset m_{ws}	1300.00	688.00	100.00	688.00

of the bearings were taken from the paper [36]. The bogie pivot is described as a spherical connection with three rotational degrees of freedom, in which the change of friction forces was determined from:

$$F_x = (W_x \cdot \chi) / \sqrt{\left(1+\left(\frac{Wm\,s}{rN_2}\right)\wedge 2\right)},$$

$$F_y = (W_y \cdot \chi) / \sqrt{\left(1+\left(\frac{Wm\,s}{rN_2}\right)\wedge 2\right)}, \qquad (12)$$

$$F_z = (W_z \cdot \chi) / \sqrt{\left(1+\left(\frac{Wm\,s}{rN_2}\right)\wedge 2\right)},$$

where: W_x, W_y, W_z denotes torsional speed of the pivot around the axle X, Y, Z; W_m is the absolute speed between the wagon body and the bogie at the central pivot point, χ_s is the tangent at the starting point of the force/speed transfer function of $3.0 \cdot 10^6$ Ns/m, N denotes the normal force directed in the Z, μ_2 is the friction coefficient of 0.19 value, r is the radius of central pivot curvature of 0.19 m [36].

Friction forces F_{bx}, F_{bx} in the lateral bearing planes, through which the body was additionally supported on the bogie, was determined from formulas (13) [5]:

$$F_{bx} = (V_x \cdot \chi) / \sqrt{\left(1+\left(\frac{Wm\,l}{N_1}\right)\wedge 2\right)}$$

$$, \qquad (13)$$

$$F_{by} = (V_y \cdot \chi) / \sqrt{\left(1+\left(\frac{Wm\,l}{N_1}\right)\wedge 2\right)}$$

where: V_x, V_y are the relative velocities in the bearing plane in the X, Y directions, Wm denotes the total relative speed on the XY plane, χ describes the tangent at the starting point of the force/speed transfer function of $3.0 \cdot 10^6$ Ns/m, N is the force induced by the weight of the wagon body operating in the direction normal to the XY plane, μ_1 is the friction coefficient equal to 0.36 [5]. Due to the fact that the paper concerns the analysis of rail vehicle derailment, the structure of the freight wagon model and its mathematical notation are described in the paper generally in the matrix form (11). However, more attention was paid to the description of wheel-rail contact and the methodology for determining forces in the wheel-rail contact zone and derailment ratio resulting from the relations between these forces.

4.2. Wheel-rail contact model

The mathematical model of a freight wagon mounted on Y25 series bogies has been integrated with algorithms and numerical procedures determining the wheel-rail contact. Numerical procedures regarding wheel-rail contact were used to determine the value of forces and areas of their action in contact zones. The contact model was based on Kalker's simplified theory [22] and the FASTSIM algorithm [23]. In order to calculate the tangential contact forces, normal pressure forces were determined, the friction coefficient was adopted at the level of 0.36 [5], the length of the axis a and b of the ellipse of the contact area (Fig. 11a) were calculated using Hertz's theory [19]. Creepage values were considered as relative stiff slip. Relations between these parameters are described by the formula:

$$\begin{matrix} rs_x \\ rs_y \\ rs_z \end{matrix} = \frac{1}{vu_x}\begin{bmatrix} sv_x \\ sv_y \\ sv_y \cdot \sin(|\alpha|) + sv_z \cdot \cos(|\alpha|) \end{bmatrix}, \qquad (14)$$

where: rs_x, rs_y are relative stiff slip/creep in longitudinal X and lateral Y directions, rs_z denotes spin, ω describes the contact angle, vu is the speed of the moving reference system, the value of which is equal to the speed of the vehicle, sv denotes the slip speed. Slip speeds sv projected into a given direction were determined from the formula:

$$\begin{matrix} sv_x \\ sv_y \\ sv_z \end{matrix} = \begin{bmatrix} vu_x \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ vr_y \\ 0 \end{bmatrix} + \bar{\omega}_w \times \bar{r}, \qquad (15)$$

where: $\bar{\omega}_w$ denotes relative $\bar{r}$ angular velocity of the wheel, $\bar{r}$ the coordinate of the contact point in the reference system connected to the centre of wheel mass, vr denotes the relative speed of the centre of wheel mass as defined in the moving reference system.

The next step was to determine tangential contact forces T_x and T_y using the FASTSIM procedure [23]. The algorithm of this procedure divides the elliptical contact zone into smaller areas/cells. In each of these cells, shear stresses and micro-slips are determined in two directions, longitudinal v_x and lateral v_y to the direction of the vehicle ride (Fig. 11b). Then, the contact zones in the contact area at $\gamma = 0$ are determined. The parameters of contact geometry as input quantities for the FASTSIM procedure have been tabulated depending on the lateral displacement of a particular wheel (Fig. 12 and 13), thus the computational time during the simulation of vehicle motion dynamics was reduced.

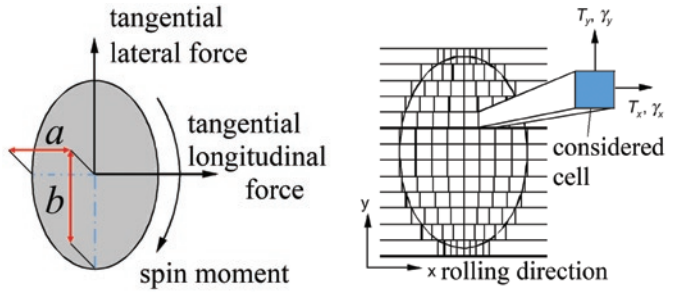
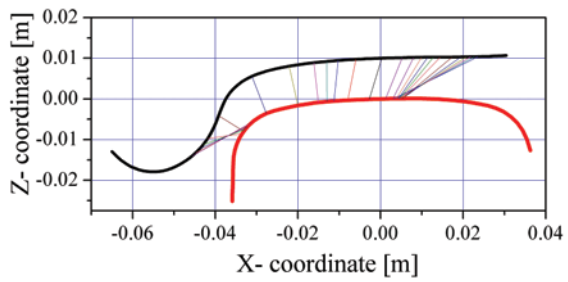


Fig. 11. Tangential forces in the wheel-rail contact zone, half-axes a and b of the contact ellipse and the division of its surface into elements



13. Wheel rail contact position for the wheel profile UIC 60 and rail profile 49E1

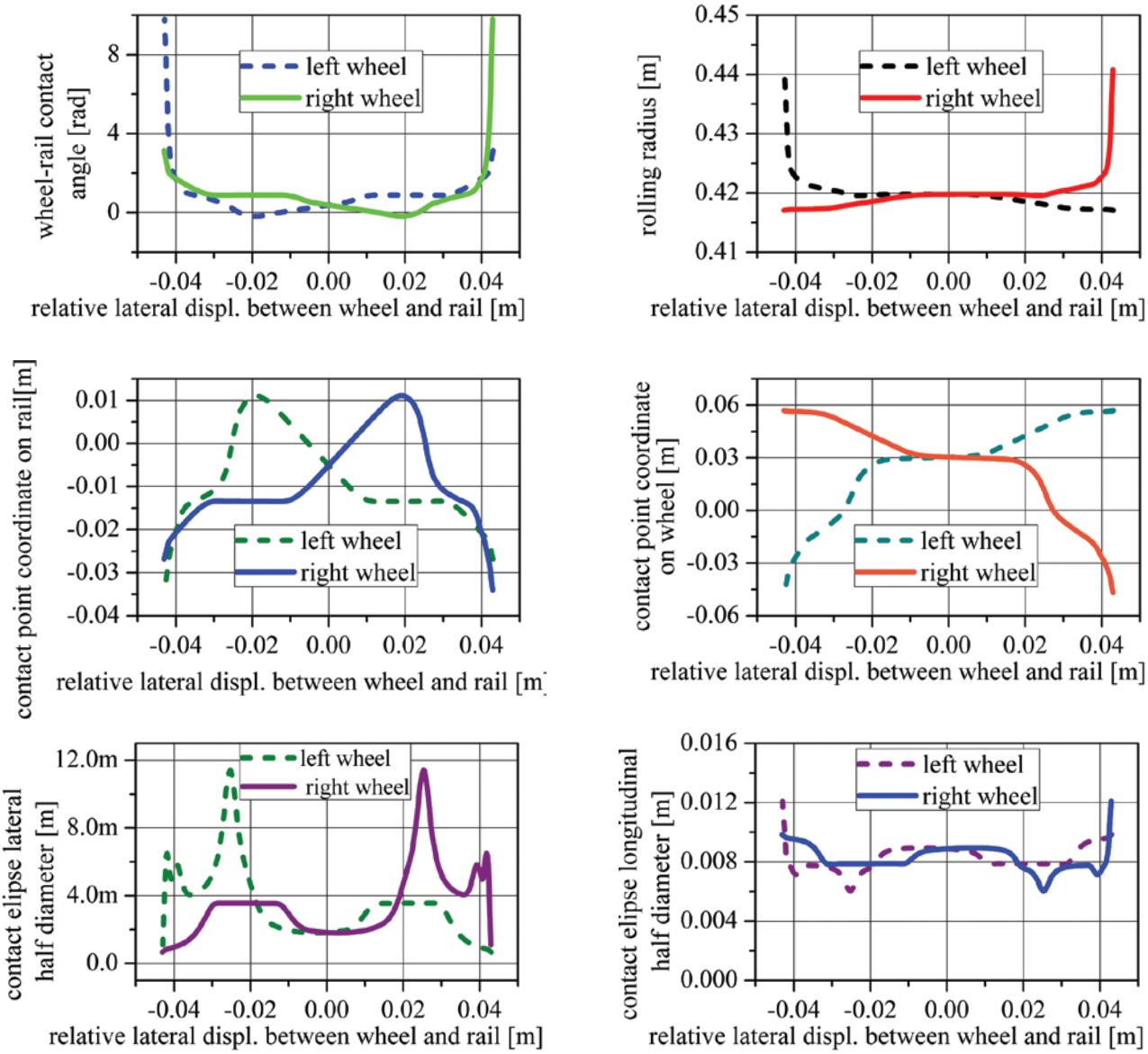


Fig. 12. Parameters of contact geometry of a UIC 60 wheel with a 49E1 rail in the function of lateral displacement y of the wheelset

The further stage of the calculation procedure involves determining and calculating the right sides of the dynamic motion equations (11) for the respective vehicle components and preparing for further solution by numerical integration of these motion equations describing the rail-track system under consideration. The approach described in this chapter allowed determining the values of forces in the wheel-rail contact zones, which were further used to calculate the value of the *Y/Q* derailment quotient.

The real nominal profiles of UIC 60 railway wheels and 49E1 rail [38], in the rail inclination configuration in the 1:40 track and shown

in Figure 14 were used in the simulation model for testing the dynamics of the freight wagon.

4.3 Results of numerical investigations

The scope of analyses carried out using the above-mentioned freight wagon-track model relates to the dynamics of the vehicle moving along a curved S-type track with a radius of R=150 m and a twist of 3 ‰ (Fig. 12). The track twist enabling tests for the vehicle and the bogie to be determined during a single ride by introducing an additional uplift in the vertical track profile (Fig. 12) was adopted as the original work input for the tests. The most unfavourable variant of the vehicle's operational configuration was adopted during these tests, i.e. the vehicle without a load. During the ride, vertical wheel forces Q and lateral (lateral to the direction of ride) Y in the wheel/rail contact zone are to be recorded. Based on the forces determined in the contact zone, Y/Q derailment quotient is to be determined as the maximum value of the lateral force to the vertical force. If the value of the Y/Q quotient does not exceed 1.2, the criterion of derailment risk is met. This criterion is based on the balance of forces in the inclined plane of the wheel/rail contact, and its limit value for a given wheel profile and adopted friction coefficient is determined from the formula:

$$\frac{|Y|}{|Q|} < \frac{tg70 - 0.36}{1 + 0.36 \cdot tg70}. \tag{15}$$

In the case when the value obtained during dynamic tests is Y/Q> 1.2, then the safety against derailment should be additionally checked by the wheel climb. This climb from the zero position should not exceed Δz=5 mm. During the numerical simulation, a vehicle travel speed of v= 5 km / h was assumed, in accordance with experimental tests (Method 3).

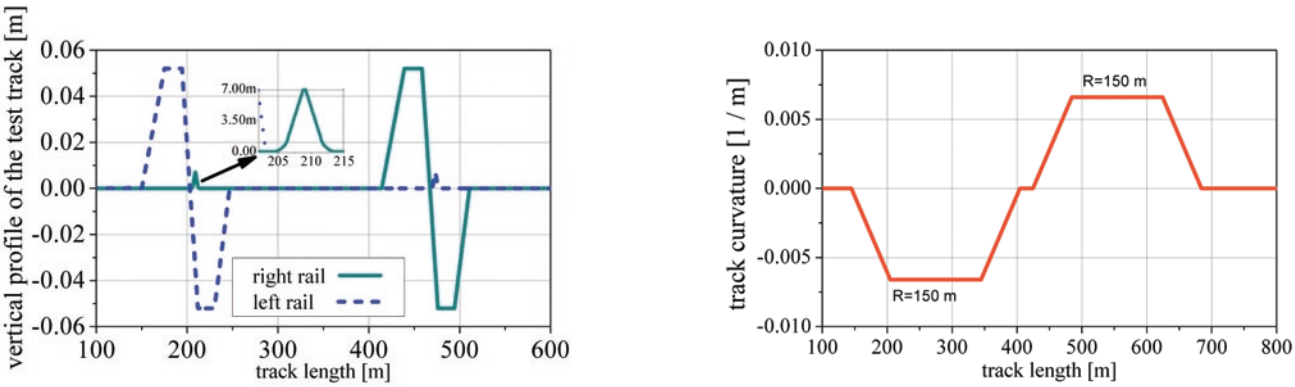


Fig. 14. Geometry of the modified test track with additional twist/lift for a wheelset test

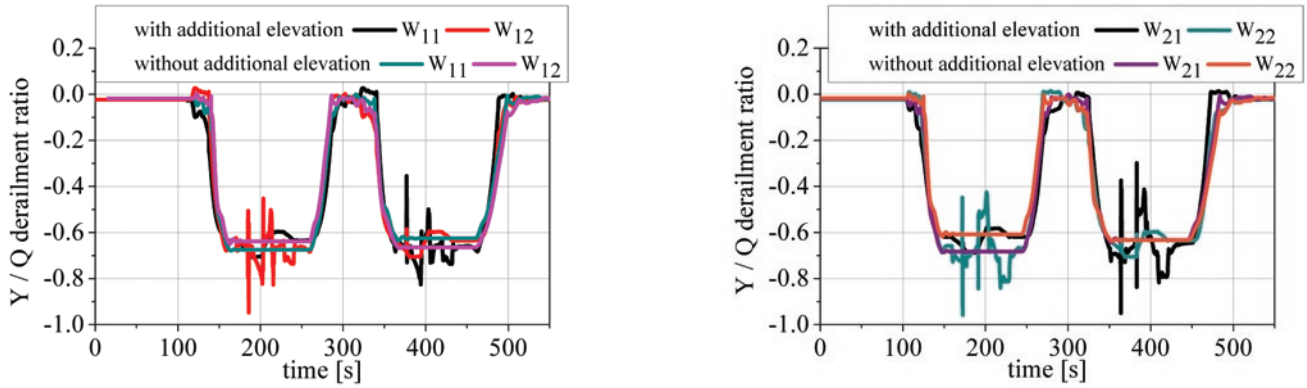


Fig. 15. Derailment quotient obtained from running over a modified test track, guiding wheelset of the first bogie W_{11} and W_{12} and the second wheelest W_{21} and W_{22}

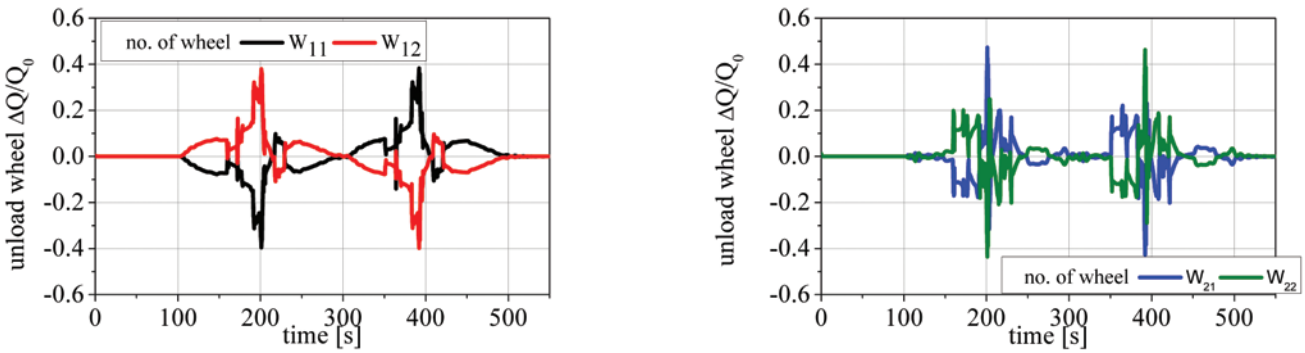


Fig. 16. Unloading the wheels of the wheelset guiding the first W_{11} and W_{12} and the second wheelest W_{21} and W_{22} of the same bogie during the passage along the modified test track

Table 4. Derailment quotient maximum values obtained from numerical investigations and the module of percentage error relative to experimental tests results

Wheel W_{ij}	11	12	21	22	31	32	41	42
Y/Q [kN]	0.67	0.63	0.23	0.27	0.56	0.61	0.68	0.66
Relative error [%]	7.89	12.76	9.52	12.50	5.48	11.46	3.03	6.45

The courses of derailment quotients shown in the graphs (Fig. 15) illustrate the comparison of results of track runs without an additional elevation (dashed curves) and with the elevation corresponding to the track twist based on the wheelbase of a single bogie. It can be observed that the values of the *Y/Q* quotient for the track without the additional lift have values convergent with the results obtained experimentally (Table 2). In order to verify the correctness of the model operation, the results of the maximum values of the *Y/Q* derailment quotient obtained during the track run without additional uplift are presented in Table 4. This allowed comparing the experimental tests with the numerical investigations, which results in establishing the module of the relative percentage error between them. The value of this error for individual wheels ranged 3.03÷12.76%. For wheels on the outer rail track side, the error between results is greater, but does not exceed 13%. The obtained result proves a correctly formulated description of the freight wagon model features and is the basis for establishing correct model validation. A much better quantitative compliance was obtained while determining the resistance yaw torque by comparing the results from numerical investigations and those obtained from the experimental tests described in Chapter 3 (Fig. 4). The created simulation model of the wagon-track system also allows to determine the unloading of the tested vehicle wheels, which is shown in Figures 16, which are also a key indicator of operational safety verification against the derailment.

5. Conclusions and summary

The article features the phenomenon of rail vehicle derailments. Methods of assessing the risk of derailment have been classified basing on a review of literature and standards. The presented description of experimental running safety tests against derailment showed the methodologies used for their implementation. Obtained and developed results from experimental tests carried out according to the above-described research methods allowed validating the results gained from the numerical investigations of the prepared model. Tests determining the course of derailment quotients' values, wheel unloading and torques were based on a numerical model describing the experimentally tested freight wagon. The paper also proposes a new, innovative test track geometry for testing safety against derailment. This innovation consisted in the introduction of an additional vertical rise in the rail track acting as twists (Fig. 14), which is based on the vehicle wheel base in a single bogie. The advantage of using in reality such track geometry during real tests, including numerical investigations, may allow reducing the costs of vehicle tests and tests of placing to service by shortening the time of such tests.

This saving results from the fact that three rides were carried out on such a modified real track, without the need for additional tests according to other risk and safety assessment methods against derailment described in Chapter 2 of this paper. The numerical results obtained using a non-linear railway vehicle model demonstrate how operational safety is influenced by various factors related to vehicle and track construction. Based on the conducted research, it has been pointed out that the *Y/Q* quotient strongly dependents on the twist resulting from the bogie base. The analysis of the results obtained from numerical investigations did not indicate a risk of derailment even with a modified track, which was also confirmed by experimental tests carried out according to the described normative methods. An error estimated below 13% between the experimental and theoretical results of the derailment ratio shows high compatibility of the theoretical model with a real rail vehicle. It should also be noted that the difference of the wheelsets angles of attack of a single bogie and the rotation of the bogie relative to the vertical axis has a significant impact on the repeatability of the results achieved.

References

1. Bogacz R, Czyczuła W, Konowrocki R. Influence of sleepers shape and configuration on track-train dynamics, Shock and Vibration 2014; Article ID 393867-1-7: 1-7, https://doi.org/10.1155/2014/393867.
2. Bogacz R, Frischmuth K. On dynamic effects of wheel-rail interaction in the case of Polygonalisation, Mechanical Systems and Signal Processing 2016; 79: 166-173, https://doi.org/10.1016/j.ymssp.2016.03.001.
3. Bogacz R. Konowrocki R. On new effects of wheel-rail interaction, Archive of Applied Mechanics 2012; 82: 1313-1323, https://doi.org/10.1007/s00419-012-0677-6.
4. Bogdevicius M, Zygiene R. Research of dynamic processes of the system "vehicle - track" using the new method of vehicle wheel with metal scale. Eksploatacja i niezawodnosc - Maintenance and Reliability 2018; 20 (4): 638-649, https://doi.org/10.17531/ein.2018.4.15.
5. Bosso N, Gugliotta A, Somà A. Simulation of a freight bogie with friction damper, Politecnico di Torino, 5th ADAMS/Rail users conference, Harlem, The Netherlands - May 2000.
6. Chudzikiewicz A, Bogacz R, Kostrzewski M, Konowrocki R. Condition monitoring of railway track systems by using acceleration signals on wheelset axle-boxes, Transport 2018; 33; 2: 555-566, https://doi.org/10.3846/16484142.2017.1342101.
7. Chudzikiewicz A, Opala M. Application of computer simulation methods for running safety assessment of railway vehicles in example of freight cars, Applied Mechanics and Materials 2008; 9: 61-69, https://doi.org/10.4028/www.scientific.net/AMM.9.61.
8. Dailydka S, Lingaitis LP, Myamlin S, Prichodko V. Mathematical model of spatial fluctuations of passenger wagon. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2008; 4 (40): 4-8.
9. Dyniewicz B, Bajer CI, Matej J. Mass splitting of train wheels in the numerical analysis of high speed train-track interactions, Vehicle System Dynamics 2015; 53(1): 51-67, https://doi.org/10.1080/00423114.2014.982659.
10. Elkins J, Wu H. New criteria for flange climb derailment, Proceedings of the ASME/IEEE Joint Railroad Conference 2000; 1-7.
11. EN 14363, Railway applications - Testing for the acceptance of running characteristics of railway vehicles - Testing of running behaviour

and stationary tests, European Committee For Standardization, 2016.
12. Federal Railroad Administration. Track Safety Standards, Part 213. Subpart G. September, 1998.
13. Garcia de Jalon J, Bayo E. Kinematic and Dynamic Simulation of Multibody Systems. Springer-Verlag, 1994, https://doi.org/10.1007/978-1-4612-2600-0.
14. Garg VK, Dukkipati RV. Dynamics of Railway Vehicle Systems. Academic Press 1984, https://doi.org/10.1016/B978-0-12-275950-5.50014-6.
15. Gaspar P, Szabo Z, Bokor J. Observer based estimation of the wheel-rail friction coefficien. IEEE Conference on Computer Aided Control System Design 2006: 1043-1048, https://doi.org/10.1109/CCA.2006.285990.
16. Ge X, Wang K, Guo L, Yang M, Lv K, Zhai W. Investigation on derailment of empty wagons of long freight train during dynamic braking. Shock and Vibration 2018; 18 Article ID 2862143, https://doi.org/10.1155/2018/2862143
17. GosNIIV-VNIIZhT: Norms for analysis and design of railway wagons MPS 1520 mm (not self-propelled). 1996.
18. He J, Ben-Gera T, Liu X. Risk analysis of freight-train derailment caused by track geometry defect. ASME. ASME/IEEE Joint Rail Conference, 2016 Joint Rail Conference, https://doi.org/10.1115/JRC2016-5743.
19. Hertz H. Über die berührung fester elastischer Körper (On the contact of rigid elastic solids). J. Reine und Angewandte Mathematik 1882; 92: 156-171, https://doi.org/10.1515/crll.1882.92.156.
20. Iwnicki S. (ed.) Handbook of Railway Vehicle Dynamics, CRC Press Inc., 2006, https://doi.org/10.1201/9781420004892.
21. Flammini F. Railway safety, reliability, and security: Technologies and Systems Engineering, IGI Global, 2012, https://doi.org/10.4018/978-1-4666-1643-1.
22. Kalker JJ. Three-dimensional elastic bodies in rolling contact. Springer, 1990, https://doi.org/10.1007/978-94-015-7889-9.
23. Kalker JJ. A fast algorithm for the simplified theory of rolling contact, Vehicle System Dynamics 2007;11 (1): 1-13, https://doi.org/10.1080/00423118208968684.
24. Kardas-Cinal E. Spectral distribution of derailment coefficient in non-linear model of railway vehicle-track system with random track irregularities. ASME. J. Comput. Nonlinear Dynam. 2013;8(3):031014-031014-9, https://doi.org/10.1115/1.4023352.
25. Koci HH, Swenson C A. Locomotive wheel-loading a system approach. General motors electromotive division. LaGrange, IL, February, 1978.
26. Konowrocki R, Bajer CI. Friction rolling with lateral slip in rail vehicles, Journal of Theoretical And Applied Mechanics 2009; 47(2): 275-293.
27. Krishna, VV, Berg M, Stiche, S. Tolerable longitudinal forces for freight trains in tight S-curves using three-dimensional multi-body simulations. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 2019, https://doi.org/10.1177/0954409719841794.
28. Krzyżyński T. On continuous subsystem modelling in the dynamic interaction problem of a train-track-system, Vehicle System Dynamics 2007; 24:311-324, https://doi.org/10.1080/00423119508969633.
29. Liu X, Saat MR. Barkan Ch. Freight-train derailment rates for railroad safety and risk analysis. Accident Analysis & Prevention 2017; 98: 1-9, https://doi.org/10.1016/j.aap.2016.09.012.
30. Matej J. Controlled wheel flange climb derailment of the load-measuring wheel set on curved and straight track, Proceedings of the Institute of Vehicles 2015;1(101): 75-90.
31. Matej J. A new mathematical model of the behaviour of a four-axle freight wagon with UIC single-link suspension. Proceedings of the Institution of Mechanical Engineers Part F-Journal of Rail and Rapid Transit 2011; 225(6): 637-647, https://doi.org/10.1177/0954409711398173.
32. Matsudaira T. Dynamics of high speed rolling stock, Japanese National Railways RTRI Quarterly Reports, Special Issue, 1963.
33. Matsumoto A, et al. Continuous observation of wheel/rail contact forces in curved track and theoretical considerations, Vehicle System Dynamics 2019; 50(1): 349-364, https://doi.org/10.1080/00423114.2012.669130.
34. Miwa M, Oyama T. Modeling an optimal track maintenance and management strategy in consideration of train derailment accident risk, Journal of Japan Society of Civil Engineers 2019; 75(1): 11-28, https://doi.org/10.2208/jscejipm.75.11.
35. Okamoto I, Uchida M. The coefficient of friction in railway vehicles and tracks. Journal of Japanese Society of Tribologists 2002; 47(4): 249-254.
36. Opala M. Evaluation of bogie centre bowl friction models in the context of safety against derailment simulation predictions, Archive of Applied Mechanics 2018; 88(6): 943-953, https://doi.org/10.1007/s00419-018-1351-4.
37. Piotrowski J, Październiak P. Influence of dither generated by rolling contact on friction damping in freight wagons, Vehicle System Dynamics 2010; 48: 195-209, https://doi.org/10.1080/00423111003706722.
38. Polish State Railways, Inc.: Instructions Id-1 (D1) - Technical conditions for the maintenance of the surface of the railway lines. Warsaw 2005.
39. Rausand M. Risk Assessment: Theory, Methods, and Applications, Wiley, 2011, https://doi.org/10.1002/9781118281116.
40. Raport ORE/ERRI B55 Rp.8 - Prevention of derailment of goods wagon on distorted tracks, 1983.
41. Reliability, Safety, and Security of Railway Systems. Modelling, Analysis, Verification, and Certification, Fantechi A., Lecomte T. Romanovsky A. (Eds.) Second International Conference, RSSRail 2017, Pistoia, Italy, November 14-16, 2017, Proceedings Series Springer Volume 10598, 2017.
42. Riessbeger K. Zur entgleisungssicherheit der rollenden landstrasse. ZEVRail Glasers Anna-len. No 2/3 1994.
43. Sato E, et al., Lateral force during curve negotiation of forced steering bogies, Quarterly Report of RTRI 2003; 44(1): 8-14, https://doi.org/10.2219/rtriqr.44.8.
44. Shabana A. Dynamics of multibody systems. Cambridge University Press, Third Editon, 2005, https://doi.org/10.1017/CBO9780511610523.
45. Szolc, T. Simulation of dynamic interaction between the railway bogie and the track in the medium frequency range. Multibody System Dynamics 2001;6(2): 99-122, https://doi.org/10.1023/A:1017513021811.
46. UIC 518, 2009, Testing and approval of railway vehicles from the point of view of their dynamic behaviour - Safety - Track fatigue - Ride quality.
47. VII001, AAR Mechanical Division, Manual of Standards and Recommended Practices. Section C- Part II, Volume 1, Chapter XI. Section 11.5.2 Track-Worthiness Criteria, Adopted 1987, Revised 1993.

48. Weinstock H. Wheel climb derailment criteria for evaluation of rail vehicle safety, Paper no. 84-WA/RT-1, 1984 ASME Winter Annual Meeting, New Orleans, LA, November, 1984.
49. Wu H, Wilson, N. Railway vehicle derailment and prevention. In: Iwnicki, S. ed. Handbook of Railway Vehicle Dynamics, Chapter 8, Taylor & Francis, Boca Raton, FL, 2006, https://doi.org/10.1201/9780849333217.ch8.
50. Zhou H, Zhang J, Hecht M. Three-dimensional derailment analysis of crashed freight trains. Vehicle System Dynamics 2014; 52(3): 341-361, https://doi.org/10.1080/00423114.2014.881512.

Robert KONOWROCKI
Institute of Fundamental Technological Research, Polish Academy of Sciences
ul. Pawińskiego 5B, 02-106 Warsaw, Poland

Railway Institute
ul. Chłopickiego 50, 04-275 Warsaw, Poland

Andrzej CHOJNACKI
Railway Institute
ul. Chłopickiego 50, 04-275 Warsaw, Poland

E-mails: rkonow@ippt.pan.pl, achojnacki@ikolej.pl

Article citation info:
WANG Z-B, LI W-Y, SHANG S, WANG Z, HAN C-Y. Performance degradation comparisons and failure mechanism of silver metal oxide contact materials in relays application by simulation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 86–93, http://dx.doi.org/10.17531/ein.2020.1.10.

ZhaoBin WANG
WeiYan LI
Shang SHANG
Zhan WANG
ChunYang HAN

PERFORMANCE DEGRADATION COMPARISONS AND FAILURE MECHANISM OF SILVER METAL OXIDE CONTACT MATERIALS IN RELAYS APPLICATION BY SIMULATION

PORÓWNANIE OBNIŻENIA CHARAKTERYSTYK ORAZ BADANIE MECHANIZMU USZKODZEŃ MATERIAŁÓW STYKOWYCH Z KOMPOZYTÓW TYPU SREBRO-TLENEK METALU STOSOWANYCH W PRZEKAŹNIKACH ELEKTROMAGNETYCZNYCH NA PODSTAWIE DANYCH Z SYMULACJI KOMPUTEROWEJ

To evaluate the electrical contact behaviors of silver metal oxide contact materials in relays application more accurately, and to guide the selection of contact materials, the test device and testing method for simulating electrical contact performance in relays application were analyzed in this paper. The electrical contact simulation test system was designed and developed, which can easily simulate contact materials. The contact resistance, static force and rebound energy degradation parameters of $AgSnO_2$, AgCdO and AgNi contact materials under the same load conditions were obtained through experimental research, the contact resistance and arcing energy degradation parameters of $AgSnO_2$ under different opening distances were acquired at the same time. The result indicated that accurate data are received by the electrical contact simulation testing method. Finally, based on the test data, the degradation performance of three selected test materials was tested, and the failure mechanism of $AgSnO_2$ materials was analyzed.

Keywords: *electrical contact performance; contact material; degradation parameter; contact resistances; failure mechanism.*

W celu dokładniejszej oceny zachowania styków elektrycznych z kompozytów srebra i tlenku metalu stosowanych w przekaźnikach elektromagnetycznych oraz w celu ułatwienia wyboru materiałów stykowych, w niniejszej pracy przeanalizowano urządzenie testowe oraz metodę testowania, które pozwalają na symulację działania styku przekaźnika. Zaprojektowano i zbudowano system testowania styków elektrycznych, który umożliwia łatwą symulację zachowania materiałów stykowych. Parametry degradacji rezystancji zestykowej, siły statycznej oraz energii odbicia materiałów stykowych $AgSnO_2$, AgCdO i AgNi uzyskano w badaniach eksperymentalnych prowadzonych w takich samych warunkach obciążenia. Jednocześnie badano także parametry degradacji rezystancji zestykowej energii łuku $AgSnO_2$ przy różnych odległościach otwarcia styków. Wyniki pokazują, że proponowana metoda badania symulacyjnego styków elektrycznych pozwala na uzyskanie dokładnych danych. W oparciu o dane testowe, przebadano zachowanie degradacyjne trzech wybranych materiałów oraz przeanalizowano mechanizm uszkodzenia styków z kompozytu $AgSnO_2$.

Słowa kluczowe: *wydajność styku elektrycznego; materiał stykowy; parametr degradacji; rezystancja zestykowa; mechanizm uszkodzenia.*

1. Introduction

The electrical properties of the contact material play a key role in the reliability of the automation system. Generally speaking, the performance indicators of the contact materials include contact resistance, resistance to fusion welding, corrosion resistance, hardness, strength and other related parameters [19]. Qualified contact materials should have good electrical and thermal conductivity, and should also have good mechanical, chemical, and arc-resistant properties as well [18]. In the working process of the contacts of the switch apparatus, there are complex mechanical, electrical, thermal and arc energy conversions, so the traditional analysis method cannot effectively analyze the failure mechanism of low-voltage electrical appliances[10]. Because the silver-based contact material has good conductivity and thermal conductivity and has high oxidation resistance and small contact resistance, it is widely used in the field of low-voltage relays application [15]. $AgSnO_2$, AgCdO and AgNi composite contact materials are widely used for controlling products [9].

It is well known that the electrical life of a low-voltage switching devices is directly related to factors such as switching load conditions,

mechanical motion characteristics, working environment atmosphere, and contact material composition[2]. Therefore, the correspondence between the material composition of a single contact material and the electrical life of the switch is not deterministic. The characteristics of the variety of electrical types also require that we should discuss the relationship respectively among the arc of the contact material, the electrical contact performance and the electrical life of the switch. The relationship between the contact material and the electrical life of the switch has been drawing much attention from relay or material manufacturers. Therefore, choosing which kind of experimental simulation to fulfil an authentic and accurate evaluation has become a critical issue.

The contact fusion welding, excessive contact resistance and contact electrical erosion are three typical failure modes that occur during the electrical life test of relays [3]. The electric life test systems currently used can be classified into two types. One is to take the motor or exciter as the driving mode to simulate the contact pair splitting process [7]. Because the speed of the simulation is in the range of a few millimetres per second, the arc characteristic parameters can be captured during the test. If the strain force sensor is introduced, the welding force in the contact breaking process can be measured together [13].Another method is to use the actual switch product as the carrier, perform electrical life test after assembling the contacts [11], and evaluate the performance of contact materials by testing the life of the switch [16].

If the motor with a slower speed is applied, the breaking speed of the contact can achieve in the hundreds of microns per second, so that the liquid bridge and its rupture process during the breaking of the contact pair can be observed [8]. The electromagnetic coils and lever structures have been used to drive the contact pairs to split and close motion [3], and the simulated speed is within the range of 100 ~ 600 mm/s [1]. Moreover, piezoelectric force sensors are used to test the welding force of the contact, to evaluate the anti-welding performance of the contact materials [5]. Taking the electromagnetic relay as an example, the contact breaking process has a variable acceleration motion characteristic, with the maximum speed order of several hundred millimetres per second.

Compared with the actual relay, there are still some differences between the previous simulation test system and the actual relay, which are mainly reflected in the following aspects: (1) it is impossible to obtain the influence of the process adjustment parameters (magnetic clearance, contact opening distance, mechanical reaction force, static force, etc.) specifically to electrical switches on dynamic characteristics and the welding force; (2) with the limitation of the response speed of the force sensor ,the movement form and breaking speed of the contact in the simulated action process are different from the actual electrical contact action process. So the conclusions obtained cannot be directly applied to the optimization of structural parameters and the selection of contact materials of the actual relays.

In this article, we present the test device and testing method for simulating electrical contact behaviour in relays application. The test system was designed in this paper using a typical clapper electromagnetic relay as the contact material test carrier, and a multi-degree-of-freedom joint mechanical adjustment mechanism was designed to realize the relay magnetic gap and contact opening distance, mechanical reaction force, static force adjustable function. During the relay action, the contact static force, welding force, collision force, contact dynamic displacement/speed, coil current, contact arc voltage/arc current, and related time parameters [17] can be measured simultaneously. It can also synchronously test the collision, rebound and fusion welding at the contacts during the relay pull-in/release process. In this paper, $AgSnO_2$, AgCdO and AgNi contact materials are taken as the research objects, and the electrical contact simulation test system of contact materials is designed. The degradation of electrical contact characteristic parameters of three kinds of materials under the same load was compared, and the degradation of electrical performance parameters of $AgSnO_2$ at different opening distance was compared in the experiment. The variation of sensitive characteristic parameters of its electrical contact performance was analyzed, and its failure mechanism was preliminarily analyzed.

2. Novel test system for simulating and evaluating electrical contact characteristics of contact materials

At the present stage, there are two main ways to establish the electrical contact simulation test system for the contact materials: a) Using the electromagnetic coil or the exciter as the driving mechanism, simulating the breaking and closing process of the contact switch, and obtaining the contact welding force, etc. [6]; b) The actual switch product is used as the test carrier, and the switch apparatus is placed in the test system to test the arc characteristics and anti-welding characteristics of the switch apparatus [12]. Although the latter method can make a real analysis of the action characteristics of the actual electrical apparatus, it is limited by the structure of the electrical apparatus themselves, so it is impossible to change the mechanical parameters and accurately observe the erosion condition of the surface of the electrical equipment. In this test system, the electromagnet coil is used to drive the pushrod to act on the moving reed, and the dynamic and static contact points are used to simulate the on-off process of the electrical switch.

2.1. Main functions and technical indicators

Main functions: The system can change contact materials which were processed into rivet shapes freely to carry out electrical contact simulation test; adjust the mechanical parameters arbitrarily, including contact opening distance, pushrod position, empty travel (free travel of pusher) and super-path, etc., to assist in obtaining high-precision test data; real-time monitoring of the contact voltage, contact current, contact voltage drop and force signal in each switching/closing process; the static force, contact resistance, arc energy (breaking arc energy), rebound energy (closing arc energy) and relevant time parameters in each operation process were calculated according to the comprehensive electrical performance parameter program.

Technical index: contact voltage measurement range 0~48 V under DC condition, contact voltage accuracy 10 mV; contact current measurement range 0~20 A, contact current accuracy 10 mA; mechanical parameter adjustment mechanism adjustment accuracy 10 μm; contact resistance accuracy 0.1 mΩ; the time parameter accuracy is 1 μs.

2.2. Test system overall structure

The functional block diagram of the contact material electrical contact test system is shown in Fig. 1. The system is mainly divided into three parts: the mechanical system part, which realizes the dynamic adjustment of the mechanical parameters; the control conditioning part, including the action control module, electromagnet driving circuit, and the signal Conditioning acquisition module, to realize the drive control electromagnet action, amplify, filter and process the collected signals; the system software part uses Labview software to design the human-computer interaction interface, which can complete the functions of setting test parameters, test control, waveform display, and data management. The major structures and physical map of the test system are shown in Fig. 2.

In Fig. 2 the contact voltage and the contact voltage drop in the process of breaking and closing are respectively tested by the resistance voltage division method and the differential amplifier, the contact current is monitored by the Hall current sensor, the static force between the contacts is monitored by the force sensor, and the contact resistance is tested by the four-wire method. The original data

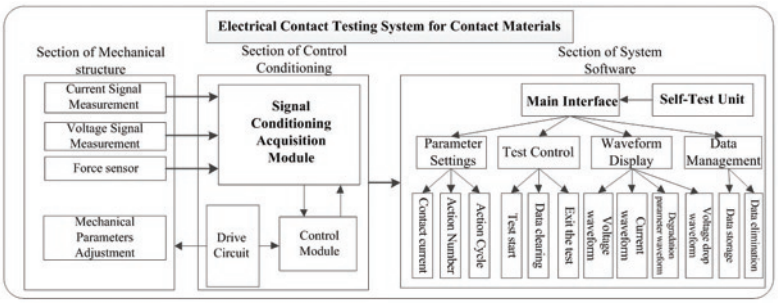
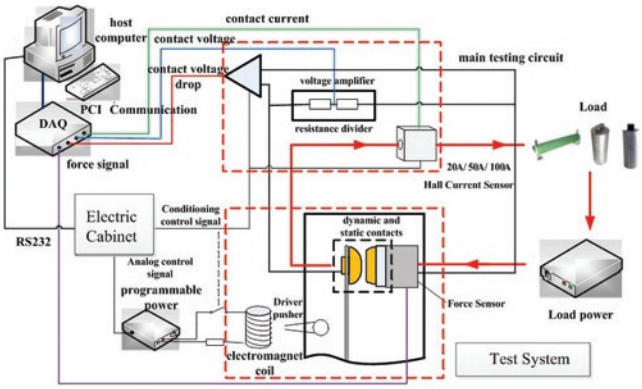
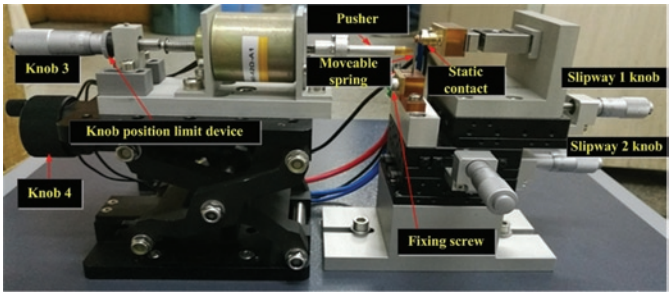


Fig. 1. Functional block diagram of the testing system for contact materials



a) Main structure of the test system



b) physical map of the test mechanical system

Fig. 2. The major structures and physical map of the test system

collected by DAQ (Data Acquisition) are sent to the upper computer through the PCI (Peripheral Component Interconnect) communication protocol, and the upper computer communicates with the lower computer through the RS232 interface to complete functions such as comprehensive parameters dynamic observation and data storage of the electrical performance simulation test. If the contact appears adhesion failure, it will prompt the tester to stop the test system and cut off the load circuit.

2.3. Mechanical structure of the test system

Mechanical structure of the simulation test system includes electromagnet drive mechanism, lifting platform, horizontal sliding table, electromagnet stroke adjustment mechanism (empty travel), pushrod point adjustment mechanism, contact opening distance adjustment mechanism, hall current sensor and static force sensor, combined with the corresponding connector to achieve the adjustment of mechanical adjustment parameters, etc. The mechanical structure is shown in Fig. 3.

The front end of the direct-acting electromagnet of the system is fixed with an insulating pushrod for simulating the movement of the electric switch pushrod, and the left side is an electromagnet stroke adjusting mechanism, and the stroke of the electromagnet can be adjusted by adjusting the differential probe. The driving mechanism and the adjusting mechanism are fixed on the Z-axis displacement slide, which can adjust the action point of the putter in the Z-axis direction of the moving reed. The dynamic and static contacts are respectively riveted on the moving reed and the static contact seat, the moving reed is fixed on the two-dimensional slide table, and the static contact seat is fixed on the two-dimensional sliding platform through the X-axis displacement sliding platform 2, the mechanical system can adjust the action point of the push rod on the Y-axis direction of the moving reed by adjusting the Y-axis displacement slide, adjust the over-travel by adjusting the X-axis displacement slide 1, and adjust the contact separation by adjusting the X-axis displacement slide 2. Adjustment of the opening distance. A static force sensor is fixed behind the static contact seat, which can detect the static force signal of the contact.

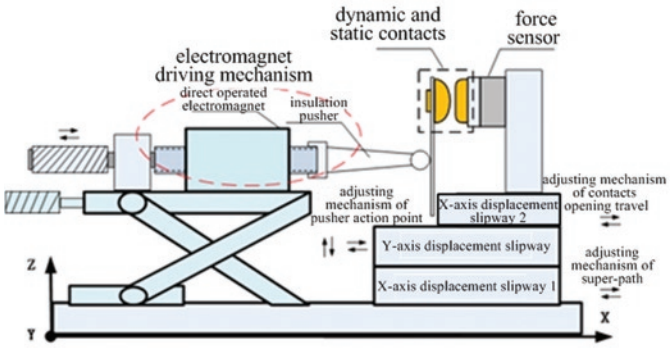


Fig. 3. Structural block diagram of mechanical system

3. Contact material performance comparison test and result analysis

3.1. Experimental test conditions

Three kinds of contact materials of $AgSnO_2$, AgNi and AgCdO were tested under the same condition. The contact test material electrical contact simulation test conditions and general composition of materials under test were shown in Table 1. The data obtained from the test were analyzed, and the variation trend curves of degradation parameters (contact resistance, static force, rebound energy) of the three materials were obtained by using the mean value method.

Table 1. Electrical contact simulation test conditions for contact materials

Test parameters	Value
Experiment environment	Room temperature: 19°C~ 22°C; Humidity: 50 ~ 65%
Contact material	$AgSnO_2$, AgNi, AgCdO
Operating frequency	1s on/1s off
Load characteristics	Pure resistive load, full process is charged
Contact voltage/current	24V/18A
Action number	Set continuous actions 20,000 times
Contact opening range	0.6mm, 0.8mm
Initial static force	1.5N

3.2. Experimental test result and analysis

3.2.1. Contact resistance

It is generally believed that the value of the contact resistance

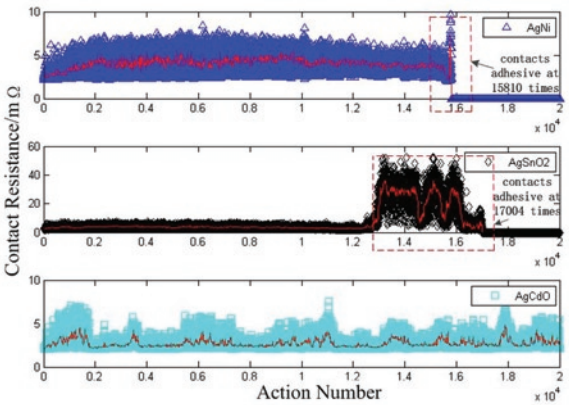


Fig. 4. Comparison of contact resistance change

depends on the resistivity, hardness, contact force and surface finish of as well as the chemical stability of the material, the magnitude of the current, conduction time, working environment and other factors [14]. In our test equipment and test conditions, the average contact resistance of $AgSnO_2$, AgNi, and AgCdO were 7.7mΩ, 4.03mΩ and 2.74mΩ respectively.

It can be seen from Fig. 4 that the contact resistance of the $AgSnO_2$ contact material tends to be stable 13,000 times before, after a sharp increase in 50mΩ situation, then the contact resistance began to fluctuate and decrease until the adhesive failure occurred at the 17004th time. In the early stage of the test, $AgSnO_2$ material has low and stable contact resistance due to arc erosion, which is not serious. With the increase of switching operations, the local molten pool will be formed on the surface of $AgSnO_2$ material under the impact of large current and arc erosion. As SnO_2 particles are hardpoints, they have good thermal stability, so no decomposition or sublimation occurs. Most of the SnO_2 particles have a lower density than the molten Ag, which will be suspended in the molten pool and form a SnO_2 particle aggregation area on the surface of the contact, resulting in increased contact resistance. In the later stage, due to the repeated thermal action of the arc, the surface topography between the contacts changes strongly, resulting in the softening of materials, the increase of metal bridge points in the fusion welding, part of the conductive spots no longer conduct electricity, and the fluctuation of contact resistance decreases until the bonding.

The contact resistance of AgCdO contact material always fluctuates around 2.5mΩ. As silver and cadmium oxide are not mutually soluble, CdO will decompose and absorb a large amount of heat under the action of arc. A part of the decomposed Cd will re-form CdO with the oxygen in the air on the surface of contact, hindering the fusion welding of moving and static contact. The contact surface of AgCdO has uniform erosion, low contact resistance, and good stability, and its resistance to fusion welding is the best.

For AgNi contact materials, under the action of 18A current, with the increase of switching operations, the erosion between dynamic and static contacts is serious and the material quality changes significantly. The pollution area increases and the contact resistance suddenly increases and a short adhesive failure occurs after 15810th time.

3.2.2. Static force

Fig.5 shows the static force test of three kinds of contact materials. The initial static force of the contact is set to 1.5N (We chose a pressure of 1.5N because the contact pressure of a certain sealed relay we analyzed was 1.5N. The test conditions are consistent with the actual relay product.). The arc of initial breaking process may belong to the anode (static contact) and the energy is high, causing the material of the initial closed contact surface to be transferred sharply. The surface of the cathode (moving contact) will have a certain convex dome, and the convex pressing of the static contact causes the interaction force between the two contacts to change, so the static force of three kinds of material maintained an upward trend at the beginning of the test. With the increase of the moving action of the moving and static contacts, the material of the contact surface accumulates, the transfer rate of the surface material slows down, and the steady trend is maintained for a long period time in the medium term.

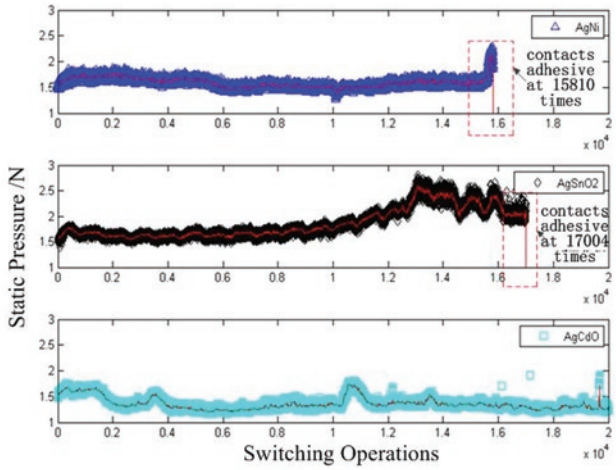


Fig. 5. Comparison of rebound energy change

In our test conditions and test equipment, the static force of the AgNi, $AgSnO_2$ contact materials fluctuate a lot before the contact fusion failure occurrs, and achieve a higher value of about 2.3N, 2.8N, respectively. This may indicate that the material transfer is large and the arc erosion is serious, dynamic contact surface microscopic spikes occurred plastic deformation, the mechanical breaking strength is not enough to make it separated, cause adhesive failure of the contact. While the static force of AgCdO increased first and then fluctuated around 1.55N all the time, and the rise of temperature was low and it's not easy to weld. Therefore, the change of static force may be directly related to the surface morphology and microstructure of the contact.

3.2.3. Rebound energy

The rebound energy (closing arc energy) and the arc energy (breaking arc energy) which are in Fig.6 and Fig.8 can be calculated from the same formula as follow:

$$Q = \int_{t_0}^{t_1} uidt = \sum_{i=1}^{n} U_i I_i \Delta t \tag{3}$$

Formulas:

- t_0 - the arcing start time (rebound start time);
- t_1 - the arcing end time (rebound end time);
- U - the contact voltage;
- I - the contact current;
- Δt - 4us in our test system (corresponds to the sampling time)
- n - $(t_1-t_0)/\Delta t$

It can be seen from Fig. 6 that the rebound energy of the three materials is maintained at 50mJ ~130mJ for most of the time, and there is a tendency to decrease at the initial stage. The reason may be that the increase of static force suppresses the vibration of the contact closure process, while the possible reason for the increase of the rebound energy before the adhesive failure of the $AgSnO_2$ contact material is the increase of the contact return time, the number of rebounds and more frequent vibration. For the AgNi and $AgSnO_2$ materials, the welding phenomenon hinders the breaking of the contacts when the bonding failure occurs, and the rebound of the contacts is suppressed to some extent, and the rebound energy shows a sharp decline trend. During the test, the rebound energy of AgCdO material kept a small fluctuation around 85mJ, and the number of contact rebounds and rebound time were relatively stable.

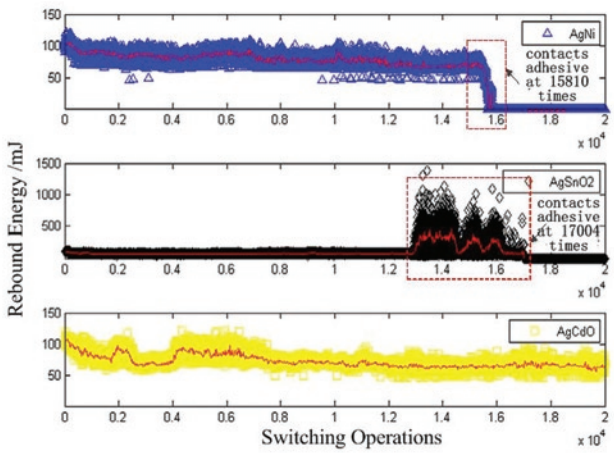


Fig. 6. Comparison of rebound energy change

Through the comparative analysis of the contact resistance, static force and rebound energy of the three materials mentioned above, it can be seen that $AgSnO_2$ has better resistance to fusion welding than AgNi, but its temperature rises higher,so adhesion is easy to occur under continuous action, and the contact resistance is higher than those of AgNi and AgCdO. AgNi contact material has poor welding resistance under slightly higher current conditions, but it has excellent processing performance with low cost, which is generally applicable to low voltage and small current microswitch appliances. In our test conditions and test device, AgCdO has better stability and strong resistance to arc ablation. However, under the continuous action of high-temperature arc, CdO will decompose Cd which is harmful to human body, so it is gradually replaced by other electric materials. The arc erosion mechanisms of the three materials are different. AgNi is the dissolution and precipitation effect derived from the metallurgical effect, while $AgSnO_2$ and AgCdO are mainly determined by the surface kinetics characteristics. In the following study, the surface morphology characteristics after material erosion will be observed through a microscope to further investigate.

4. Analysis of influence of opening distance on contact characteristics of contact materials

In our test equipment and test conditions, in order to study the influence of opening distance on the parameters of contact materials, $AgSnO_2$ material was taken as an example to make a comparative analysis on the parameters of contact resistance, arc ignition energy, arc ignition time and rebound energy under two opening distances. Because of the limited time, the relevant tests were only repeated three times, and the trend was the same. In future research work, we will continue to test and analyze the results of the experiments and the thermal effects will be considered in the future publications.

4.1. Measured results of contact resistance at two opening distances

The experimental results of contact resistance at two opening distances are shown in Fig.7.

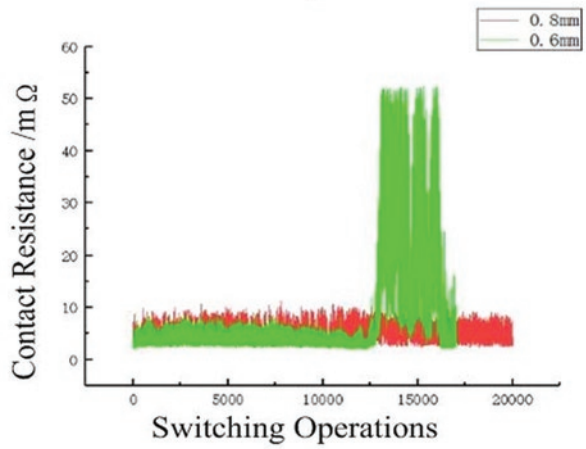


Fig. 7. Contact resistance comparison of AgSnO2 at two opening distances

Under the two opening distance conditions, the contact resistance is kept at a low level in the first and middle stages of the operation, no more than 10mΩ. In the case of an opening distance of 0.6mm, the contact resistance begins to decrease at approximately 10,000 operations. There is no significant change in the contact resistance at the 0.8mm opening distance.

4.2. Measured results of arc time and arc energy at two opening distances

The measured results of arc time (the time between first and last arc) and arc energy at two opening distances are shown in Fig.8 and Fig.9.

The waveform of arc time and arc energy are very similar. The arc energy and arc time at 0.8mm opening distance are close to the data of 0.6mm opening distance before the 1000th movements. After the 1000th movements, the arc energy and arc time are kept low at 0.8mm opening distance. The arc energy is around 100~300mJ. At the same time, after the 12,000th movements, a strong arc of more than 2000mJ begins to appear, and the frequency of strong arc increases gradually.

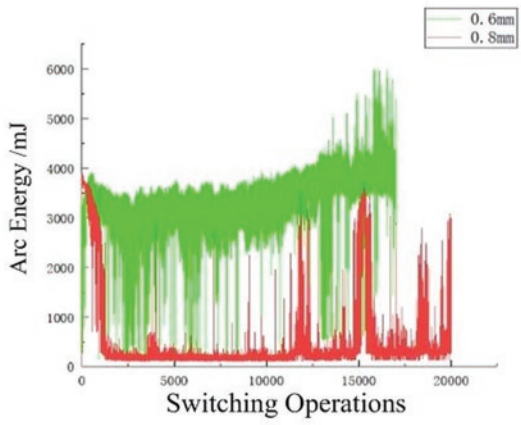


Fig. 8. Arc energy comparison of $AgSnO_2$ at two opening distances

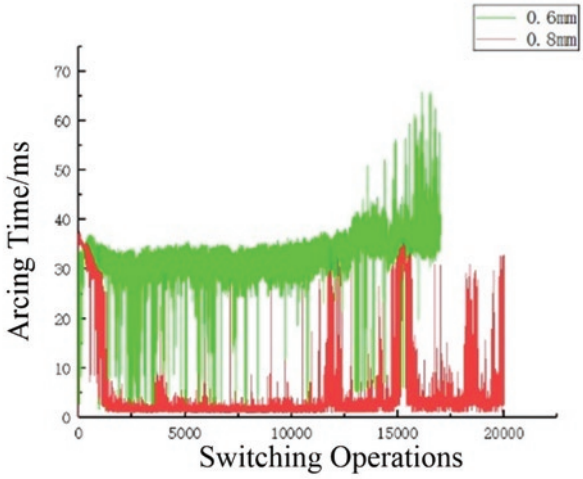


Fig. 9. Arc time comparison of AgSnO2 at two opening distances

4.3. Measured results of rebound energy at two opening distance

The measured results of rebound energy at two opening distance are shown in Fig.10.

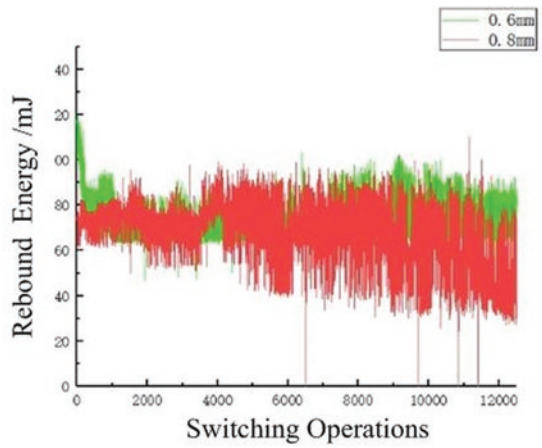


Fig. 10. Rebound energy comparison of $AgSnO_2$ at two opening distances

In the case of the opening distance d=0.6mm and d=0.8mm under our test equipment, the mean value of the rebound energy of the contact is roughly equal, and the fluctuation of the rebound energy under the opening distance of 0.8mm is more severe. The larger the contact opening distance, the greater the initial breaking force that the contact can achieve, which enhances the ability of the contact to break the weak adhesion. At the same time, the greater the breaking force is , the greater the breaking speed of the contacts will be, which is also conducive to the extinguishing of the arc.

5. Failure process analysis of $AgSnO_2$

In the electrical life simulation experiment, the contact resistance and static force of $AgSnO_2$ at 0.6mm spacing are shown in Fig. 11. At the beginning of the contact working process, the contact surface is relatively pure, so under the heat generated by the arc, the material transfer and pollution on the contact surface are more severe, which cause the surface morphology of rapid change, and the layer of oxides carbides (It is mainly oxides, but sometimes we can also detect the carbides. Perhaps it is related to the experimental atmosphere.) is accumulated on the contact surface, making the static force rise rapidly at the beginning of the work, increase from 1.4N to approximately 1.8N. Then the upward trend of static force was restrained and began to show cyclical fluctuations.

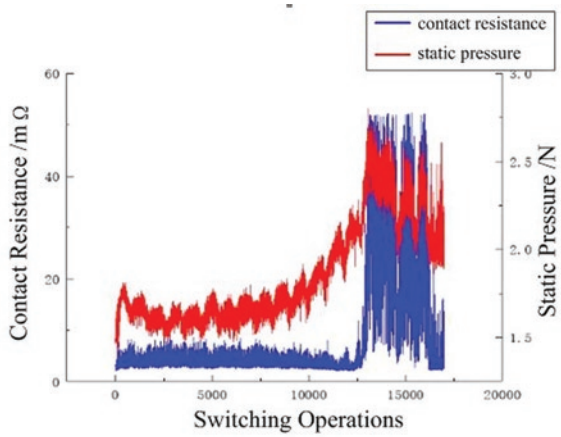


Fig. 11. Comparison of contact resistance and static force of $AgSnO_2$

According to the comparison between the contact resistance and the static force in some periods in Fig. 12, it can be found that the periodic fluctuation of the contact resistance and the static force are inversely proportional. This is because the high temperature generated by the arc causes the carbides formed on the surface to disappear during the material transfer, the contact material has a certain softening or even melting, and formed some relatively raised metal spike as the contact breaking, which reduces the actual contact area between the contacts, reduces the static force, and makes the contact resistance rise. After that, with the erosion of the arc and spatter of the material, the surface material of the contact produces loss, and some of the sharp peaks are re-worn, the actual distance between the contacts increases to the previous level, and the static force of the contact increases and while the contact resistance decreases. In the later stage of the test, with the erosion of the arc, the process mentioned above is repeated, and the contact material is continuously eroded in this process, resulting in material transfer and material loss.

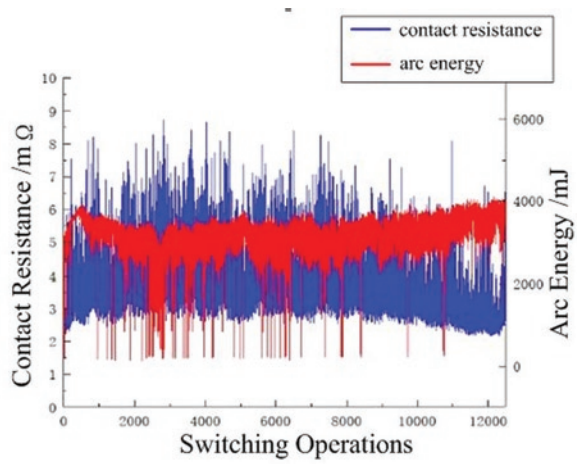


Fig. 12. Arc time comparison of $AgSnO_2$ at two opening distances

According to the contrast of contact resistance and arc energy of $AgSnO_2$ after 12000 actions in Fig. 12, the surface morphology of the contacts constantly changes and the relatively convex peaks increase due to the high temperature generated by the continuous arc, which make actual contact area between the contacts are gradually restored to the original level. Therefore, after the initial rising trend of contact resistance, it begins to decline to the initial level, and the contact re-

sistance fluctuates very sharply in this process.As the accumulates the amount of contact arc erosion, the moving contact has a large range of ridges due to material transfer, and the static contact causes additional loss due to splashing in the process of material transfer, thus it will lead to the actual distance between the contacts increases, which will aggravate the rebound phenomenon when the contacts are closed. At the same time, the metal bridge occurs between the contacts due to the welding phenomenon during the working process, which makes contact separation need larger breaking force. The static force also becomes larger, which suppresses the rebound of the contacts. The final effect of these two aspects acting on the working process of the contact at the same time is shown in Fig.13 and Fig.14.

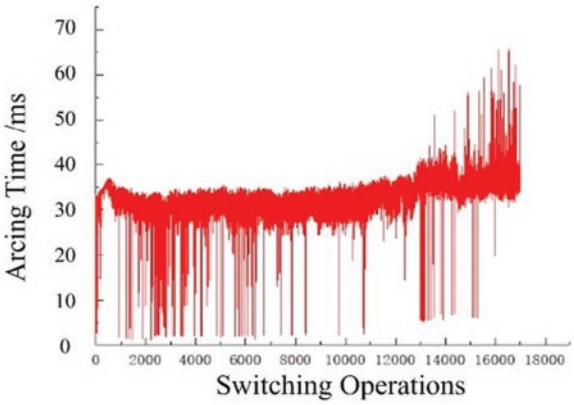


Fig. 13. Arc Burning Time of $AgSnO_2$

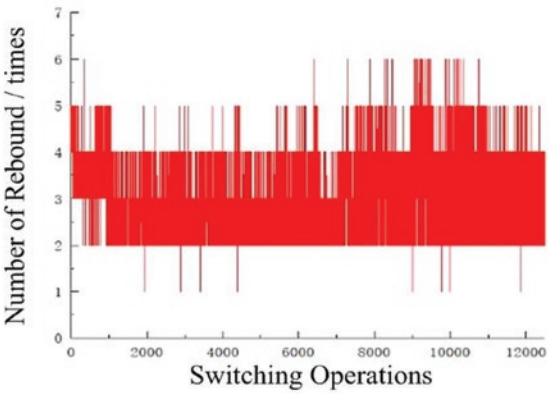


Fig. 14. Arc energy comparison of $AgSnO_2$ at two opening distances

It can be seen from the figure that, on one hand, the arc time and the arc energy increase when the contact is released, which obviously increases the amount of contact arc corrosion. On the other hand, the rebound phenomenon of the contact is slightly increased, and the frequency of 4, 5 times or even 6 times jumps increases. All these changes will aggravate contact erosion, and as this process continues, the loss of contact materials will become more and more serious, and finally there will be obvious abnormalities in its electrical properties, leading to complete wear and failure of the metal layer of contact materials.

6. Conclusion

(1) (1) An electromagnet push-rod type contact material test system is designed to simulate the action of relays. This system can analyze the influence of different contact materials, different load levels and different mechanical parameters on the electrical properties of the contact materials.

(2) By comparing the electrical properties of $AgSnO_2$ at different opening distance, the relationship between contact opening distance and electrical properties of contacts was preliminarily analyzed. It was verified that larger contact opening distance can reduce arc erosion and overcome the welding force, thus helping to extend contact life;

(3) Through the analysis of the degradation trend of electrical performance parameters during the test of $AgSnO_2$ materials, the failure mechanism is analyzed. The material mass transfer and loss caused by arc erosion, mechanical wear and other factors are the main reasons of the contact failure of the $AgSnO_2$ materials. The contact gradually develops a fusion phenomenon due to the high temperature generated by the arc, and the welding combined with the loss of the contact material causes the contact failure. While the final failure mode of the contact depends on the loss rate of the material and the severity of the fusion welding. The contact will fail at first due to the wear of the surface material if the material loss rate is fast and contact bonding failure may occur if welding is more severe.

Acknowledgment

The authors are grateful to the anonymous reviewers, and the editor, for their critical and constructive review of the manuscript. This study was co-supported by the National Natural Science Foundation of China (51507074, 61801196); Natural Science Foundation of Higher Education Institutions of Jiangsu Province (17KJB510014); Jiangsu Provincial Postgraduate Research and Practice Innovation Program Funding Project and Jiangsu Government Scholarship for Overseas Studies (JS-2018-260, JS-2018-259). Jiangsu Provincial Graduate Research and Practice Innovation Program Funding Project (KYCX19-1692).

References

1. Allen S E, Streicher E. The effect of microstructure on the eletrical performance of Ag-WC-C contact materials. Proceedings of the 44th IEEE Holm Conference on Electrical Contacts 1998: 276-285.
2. Chen Z K, Witter G. Dynamic welding of silver contacts under different mechanical bounce condition. Proceedings of 45th IEEE Holm Conference on Electrical Contacts 1999: 1-8, https://doi.org/10.1109/HOLM.1999.795920.
3. Chen Z K, Witter G. Electrical contacts for automotive applications: a review. IEICE Transactions on Electronics 2004; E87-C (8): 1248-1254.
4. Chen Z K, Witter G. The effect of silver composition and additives on switching characteristics of silver tin oxide type contacts for automotive inductive load. Proceedings of the 51st IEEE Holm Conference on Electrical Contacts 2005: 35-41.
5. Leung C, Streicher E, Fitzgerald D, et al. Contact erosion of AgSnO2In2O3 made by internal oxidation and powder metallurgy. Proceedings of the 51st IEEE Holm Conference on Electrical Contacts 2005: 22-27.
6. Liu D, Li Z B, Li C, et al. Experimental study on electromechanical characteristics of electrical contact in making and breaking operations. Low Voltage Apparatus 2013; 7(11): 10-14.

7. Liu D X, Li P Y, Rong M Z. Development of the equipment for electrical contact performance measurement. Precious Metals 2005; 26(4): 44-48.
8. Miyanaga K, Kayano Y, Komakine T, et al. effect of heat conductivity on bridge break at different material contact pairs. IEICE Transactions on Electronics 2011; E94-C(9): 1431-1434, https://doi.org/10.1587/transele.E94.C.1431.
9. Ning Y T, Zhao H Z. Silver. Changsha: Central South University Press, 2005.
10. Ren W B, Chen Y, Wang Z B, et al. Electrical contact resistance of coated spherical contacts. IEEE Transactions on Electron Devices 2016; 63(11): 4373-4379, https://doi.org/10.1109/TED.2016.2612545.
11. Ren W B, Du Y W, Man S D. A test rig for simulating the electrical performance of contact materials used in ac contactor. Electrical & Energy Management Technology 2016; (7): 11-15.
12. Ren W B, Jin J B, Li D, et al. Design and application of a novel test rig for simulating and evaluating electric arc and electrical contact characteristics of contact materials. Electrical Materials 2014; 4(7): 28-37.
13. Rong M Z, Li P Y, Li D X. The development of a new-type electrical contact performance measurement system of contacts. Electrical Engineering Materials 2005; (1): 17-21.
14. Rong M Z, Sun M. Study on surface dynamics characteristics of silver metal oxide (AgMeo) contact materials. Chinese Society for Electrical Engineering 1993; 13(6): 28-30.
15. Wang S B, Xie M, Liu M M, et al. Research progress of AgNi contact materials. Rare Metal Materials and Engineering 2013; 42(04): 875-880.
16. Wang S J, Yu Q. Study on contact failure mechanisms of accelerated life test for relay reliability. IEICE Transactions on electronics 2009; E92-C (8): 1034-1039, https://doi.org/10.1587/transele.E92.C.1034.
17. Wang S J, Yu Q, Zhai G F. The discrimination of contact failure mechanisms by analyzing the variations of time parameters for relays. IEICE Transactions on Electronics 2010; E93-C (9): 1437-1442, https://doi.org/10.1587/transele.E93.C.1437.
18. Wang, Z B, Shang S, Wang J W, et al. Accelerated storage degradation testing and failure mechanisms of aerospace electromagnetic relay. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19(4): 530-541, https://doi.org/10.17531/ein.2017.4.6.
19. Zhou L K, Man S D, Wang Z B, et al. On the relationship between contact a-spots features and electrodynamic repulsion force for electrical apparatus. IEEE Transactions on Components Packaging and Manufacturing Technology 2018; 8(11): 1888-1895, https://doi.org/10.1109/TCPMT.2018.2869993.

ZhaoBin WANG
Shang SHANG
School of Electronics and Information
Jiangsu University of Science and Technology
No. 02 Mengxi Street, Jingkou District
Zhenjiang, Jiangsu, 212003, P.R. China

School of Electronic and Electrical Engineering
University of Leeds
Leeds, LS2 9JT, United Kingdom

WeiYan LI
Zhan WANG
ChunYang HAN
School of Electronics and Information
Jiangsu University of Science and Technology
No. 02 Mengxi Street, Jingkou District
Zhenjiang, Jiangsu, 212003, P.R.China

E-mails: wangzb@just.edu.cn, 1825339708@qq.com, shangshang@just.edu.cn, 1165308663@qq.com, 605877761@qq.com, 1900562778@qq.com, 2358313772@qq.com

Article citation info:
KNEFEL T, NOWAKOWSKI J. Model-based analysis of injection process parameters in a common rail fuel supply system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 94–101, http://dx.doi.org/10.17531/ein.2020.1.11.

Tomasz KNEFEL
Jacek NOWAKOWSKI

MODEL-BASED ANALYSIS OF INJECTION PROCESS PARAMETERS IN A COMMON RAIL FUEL SUPPLY SYSTEM

MODELOWA ANALIZA PARAMETRÓW PROCESU WTRYSKU W UKŁADZIE ZASILANIA TYPU COMMON RAIL*

In this work, a simplified model of a rail-based fuel supply system in a compression ignition engine is presented. High pressure hoses were not taken into consideration and an empirical model was developed to simulate the injectors. The basic equations of the model are presented. Phenomena were described by 17 first-order ordinary differential equations. This work also contains an evaluation of the impact of the rail's geometrical parameters on the injection process. The evaluation was carried out via a program calculating the injection process, using a balance model of the injection system. A method for making preliminary decisions on the geometrical parameters of the rail is proposed.

Keywords: *common rail injection system, simulation calculations, injection process, fuel supply system operating parameters.*

W pracy przedstawiono uproszczony model zasobnikowego układu zasilania w paliwo silnika o zapłonie samoczynnym. W rozważaniach nie uwzględniono przewodów wysokiego ciśnienia, a do symulacji pracy wtryskiwaczy opracowano empiryczny podmodel. Przedstawiono podstawowe równania modelu. Zjawiska zostały opisane układem 17 równań różniczkowych zwyczajnych, pierwszego rzędu. W pracy również zawarto ocenę wpływu parametrów geometrycznych zasobnika na proces wtrysku. Ocenę przeprowadzono za pomocą programu obliczającego proces wtrysku, wykorzystującego model rozważanego układu wtryskowego. Zaproponowano sposób wstępnego doboru parametrów geometrycznych zasobnika.

Słowa kluczowe: *układ wtryskowy common rail, obliczenia symulacyjne, proces wtrysku, parametry eksploatacyjne układu zasilania*

1. Introduction

The development of contemporary high-speed compression-ignition engines is linked to the development of their fuel supply systems. Currently, common rail fuel injection systems dominate the supply of fuel to such engines. When choosing a fuel supply system for an engine, it is essential to consider many operational and control factors. Making use of simulation research for analysis of those parameters significantly eases and accelerates development work.

In the rail of a fuel supply system, the achievement of high pressure occurs in a piston-based high-pressure pump, from where a high-pressure hose leads to the fuel rail, after which relatively short hoses lead to the injectors.

Existing models describing common rail injection systems were created by research teams aiming predominantly at comparisons of their operational parameters and performance with those of other injection systems. They contained analyses and discussions of basic parameters connected to the injection event [3], yet they also allowed the determination of quantities which are difficult to measure, such as (for example) the effective cross-sectional area of the flow. Considerations can also be found dealing with compression-ignition fuel supply systems and problems associated with their control. One of the most fundamental works in this area is [2], in which the authors – as some of the first to do so – dealt with the fuel flow and control of injectors. In turn, in [4] a model of the overflow valve controlling the fuel pressure was developed, as well as a model of the fuel throttle valve in the high-pressure pump, which also considered a sub-model of that pump. The latter was controlled by a pseudo-random binary sequence. Analytical simulations of the properties of materials used in the construction of solenoid valves used in injectors and the operations of those values have also been carried out [8]. In turn, the authors of [10] focused on the stiffness of the needle-controlling rod assembly in their considerations, the values of the flow coefficients through the nozzle outlets and determination of the non-dimensional cavitation number. Determining the impact of wave phenomena in the rail on the injection event was the main topic addressed in [1], carried out by Daimler Chrysler AG, as well as [6], in which the impact of fuel properties and pressure, injection duration and the length and diameter of the injection hose on pressure changes in the system was investigated. Evaluations and selection criteria for injection systems' geometric parameters of can be found in [1, 9]. Another group of studies contains attempts to describe models aiming at control of pressure in a common rail injection system. Here, the modulus of compressibility, pressure, fuel temperature, and engine rotational speed were all taken into consideration and the results were linearized control models and initial design for controllers and regulators used for approval of systems controlling rail pressure [5, 7]. Work continues [11] on fourth-generation rail-based injection systems and on-line models adjusting the flow of fuel from the nozzle. This has served as inspiration to elaborate a simplified model of the fuel supply system.

A tendency towards the use of computer software for model ling and analysis of one-dimensional, multi-domain and mechatronic systems (interface, static and dynamic analysis) can be noted. In the majority of cases, the AMESim and Matlab/Simulink packages are used

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

– although this is not a rule. Often, initial development is conducted via those packages and more precise efforts dedicated to a specific purpose are carried out using conventional programming languages. Commonly, these are one-dimensional models, describing the non-steady, elastic flow of fuel in the system. However, simulation work is conducted based on computational models of varying degrees of complexity. Technical analyses often employ simplified models, paying attention to a reduced set of factors impacting the trace of the processes under analysis. In such cases, it is important to always determine the impact of the simplifications on the accuracy of the parameterisation of the phenomena under consideration.

After conducting initial analyses of the available literature, it was concluded that models using fluid mechanics equations models in connection with results from test benches are lacking. Thus, it was decided to undertake work on a theoretical-empirical model of the system, by means of conventional programming language.

The work presented here concerns a simplified system model, in which fuel from the three-piston high-pressure pump is fed to the inlet chamber and from there to the pressure vessel (rail).

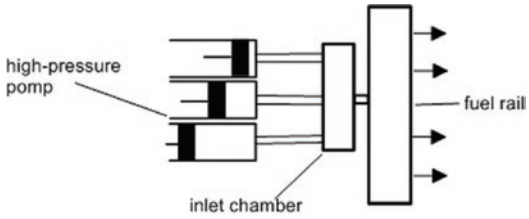


Fig. 1. Schematic of the modelled fuel supply system

In these considerations, the high pressure hose, the injector hoses and the injectors were not considered. The injection (fuel flow) proceeds through four openings, directly from the pressure vessel (Fig. 1).

2. Differential equations describing the system

Differential equations considering the system can be show in the form presented below.

2.1. Equations determining the pressure in the pump chambers

Mindful of the diminutive size of the chambers, it was taken that the change in pressure $p_p^{(i)}$ as a function of time can be can be determined from simplified continuity equations:

$$\frac{dp_p^{(i)}}{dt} = \frac{E_p^{(i)}}{V_p^{(i)}}\Big[\ A_p^{(i)}\frac{dh_p^{(i)}}{dt} - \mathrm{sgn}(p_p^{(i)} - p_d)\,\varepsilon_A^{(i)}\,\mu_d F_d^{(i)}\sqrt{\frac{2}{\rho_p^{(i)}}\left|p_p^{(i)} - p_d\right|} - \tag{1}$$

$$-\mathrm{sgn}(p_p^{(i)} - p_L)\,\varepsilon_B^{(i)}\,\mu_w F_w^{(i)}\sqrt{\frac{2}{\rho_p^{(i)}}\left|p_p^{(i)} - p_L\right|} - \varepsilon_u^{(i)} F_u^{(i)}\sqrt{\frac{2}{\rho_z^{(i)}}\left|p_p^{(i)} - p_z\right|}\ \Big]\eta_p\eta_t$$

for i = 1, 2, 3.

where: $p_p^{(i)}$ – pressure in the chamber of the ith pump piston,

$V_p^{(i)}$ – volume of the chamber of the ith pump piston,

$E_p^{(i)} = E_p^{(i)}(p_p^{(i)}, T)$ - modulus of elasticity of the fuel in the chamber of the ith pump piston,

$A_p^{(i)}$ – surface area of the ith pump piston,

$\frac{dh_p^{(i)}}{dt}$ – speed of the ith pump piston,

p_d – feed pressure,

μ_d – inlet opening flow coefficient,

$\varepsilon_A^{(i)}$ – control indicator,

$F_d^{(i)}$ – surface area of the inlet opening to the chamber of the ith pump piston,

$\rho_p^{(i)} = \rho(\rho_\rho^{(i)}, T)$ – density of the fuel in the chamber of the ith pump piston,

$\varepsilon_B^{(i)}$ – control indicator,

p_L – pressure in the inlet chamber,

μ_w – outlet opening flow coefficient,

$F_w^{(i)}$ – surface area of the outlet opening of the chamber the ith pump piston,

$\varepsilon_u^{(i)}$ – control indicator,

$F_u^{(i)}$ – surface area of the relief aperture,

p_z – rail pressure,

η_p – pump efficiency, dependent on rotational speed and fuel pressure,

η_t – correction factor considering changes in the effectiveness of the ith piston in the high-pressure pump in response to changes in the fuel temperature.

Controlling coefficients were inserted into the equations above; their interpretations are as follows:

$\varepsilon_A^{(i)}$ – volumetric outflow element, dependent on the pressure difference $p_p^{(i)}$ and p_d active only where the lift of the inlet valve $h_g^{(i)} \rangle 0$,

$\varepsilon_B^{(i)}$ – volumetric outflow element, dependent on the pressure difference $p_p^{(i)}$ and p_L active only where the lift of the valve ball connecting the pump chamber to the inlet chamber zero $h_k^{(i)}$ is greater than zero;

$\varepsilon_u^{(i)}$ – the third outflow element is active in equation (1) where the rail pressure p_z exceeds a given boundary value $p_z^{(gr)}$, simultaneously the pump piston moves upwards ($\dot{h}_p^{(i)} \rangle 0$) and the relative error of deviation of p_z from $p_z^{(gr)}$ exceeds the permissible level ε.

Moreover, it was accepted that the vent cross-sectional area $F_u^{(i)}$ changes in response to the pressure difference p_z and $p_z^{(gr)}$, according to the formula:

$$F_u^{(i)} = F_{u0}^{(i)}\sqrt{\frac{2}{\rho_z}\left|p_z - p_z^{gr}\right|} \tag{2}$$

The equations that make up (1) are non-linear first-order ordinary differential equations.

2.2. Equations of movement of the inlet valves

From Newton's second law of motion, it follows that motion of the inlet valves in a straight line is described by:

$$\frac{d^2 h_g^{(i)}}{dt^2} = \frac{\varepsilon_g^{(i)}}{m_g^{(i)}} f_g^{(i)} \tag{3}$$

where: $f_g^{(i)} = -\left(h_g^{(i)} + h_{gl0}^{(i)}\right) k_g^{(i)} - \left(p_p^{(i)} + p_d^{(i)}\right) F_g^{(i)}$

$F_g^{(i)} = \frac{\pi \left[g_1^{(i)}\right]^2}{4}$ for i = 1, 2, 3,

$h_g^{(i)}$ – inlet valve lift,

$m_g^{(i)}$ – valve mass,

$h_{gl0}^{(i)}$ – basic tension of the spring,

$k_g^{(i)}$ – spring constant,

$F_g^{(i)}$ – valve surface area.

The control indicator ε_g present in the equations above takes values:

$$\varepsilon_g^{(i)} = \begin{cases} 1 & \text{where } wG^{(i)} = 1 \\ 0 & \text{in all other cases} \end{cases}$$

In turn, indicator $wG^{(i)}$ reflects whether the valve head has returned to the valve seat ($wG^{(i)}$ = 0), or is in motion ($wG^{(i)}$ = 1), and also whether it reached its maximum lift value $h_{g_{\max}}^{(i)}$ ($wG^{(i)}$ = 2), for which $wG^{(i)} \in \{0, \ 1, \ 2\}$.

Equation (3) constitutes a system of three second-order ordinary differential equations.

2.3. Equations of the motion of the outlet ball valves

The equations of motion of the outlet valves have the following form:

$$\frac{d^2 h_k^{(i)}}{dt^2} = \frac{\varepsilon_k^{(i)}}{m_k^{(i)}} \quad f_k^{(i)} \tag{4}$$

where: $f_k^{(i)} = -\left(h_k^{(i)} + h_{kl0}^{(i)}\right) k_k^{(i)} - \left(p_p^{(i)} + p_o^{(i)}\right) F_w^{(i)}$ for i = 1, 2, 3,

$h_k^{(i)}$ – outlet valve ball lift,

$m_k^{(i)}$ – ball mass,

$h_{kl0}^{(i)}$ – initial spring tension,

$k_k^{(i)}$ – spring constant,

$F_w^{(i)}$ – surface area of the outflow aperture.

The control indicator εk present in those equations takes values:

$$\varepsilon_g^{(i)} = \begin{cases} 1 & \text{where } wG^{(i)} = 1 \\ 0 & \text{in all other cases} \end{cases} \qquad wK^{(i)} \in \{0, \ 1, \ 2\}$$

In turn, $wK^{(i)}$, as with indicator $wG^{(i)}$, indicates the position of the ball valve: 0 where no flow between the pump chamber and the inlet chamber occurs, 1 if the ball is in motion, 2 if the ball achieved its maximum lift value.

The equations of (4) constitute a system of three second-order ordinary differential equations.

2.4. Equations describing the pressure in the inlet chamber

As with the case of the high-pressure pump chamber, it was taken that changes in the inlet chamber pressure can be described by a simplified continuity equation:

$$\frac{dp_L^{(i)}}{dt} = \frac{E_L^{(i)}}{V_L^{(i)}} \Big[\ \mathrm{sgn}\,(p_p^{(i)} - p_L)\, \varepsilon_B^{(i)}\, \mu_w^{(i)} F_w^{(i)} \sqrt{\frac{2}{\rho_p^{(i)}} \left| p_p^{(i)} - p_L \right|} + \\ - \mathrm{sgn}\,(p_L - p_Z)\, \mu_L F_L \sqrt{\frac{2}{\rho_0} \left| p_L - p_Z \right|} \ \Big] \tag{5}$$

for i = 1, 2, 3,

where: $V_L^{(i)}$ – inlet chamber volume,

$\mu_w^{(i)}$ – flow coefficient for the rail inlet aperture,

F_L – area of the inflow opening to the rail, equal to the cross-sectional area of the aperture connecting the inflow chamber with the rail;

all other symbols as defined in section 2.1.

It is worth noting that volume V_L must be increased by the volume of the hose connecting the inlet chamber with the rail:

$$V_L := V_L + \frac{\pi d^2}{4} L \tag{6}$$

where: d – hose diameter, L – length of the hose connecting the inlet chamber with the rail.

2.5. Rail pressure equations

A continuity equation was used to describe the pressure differential:

$$\frac{dp_z}{dt} = \frac{E_z}{V_z} \Big[\ -\sum_{i=1}^{4} \mathrm{sgn}\,(p_z - p_k)\, \varepsilon_z^{(i)}\, \mu_z^{(i)} A_z^{(i)} \sqrt{\frac{2}{\rho_z} \left| p_z - p_k \right|} + \\ \mathrm{sgn}\,(p_L - p_Z)\, \mu_L F_L \sqrt{\frac{2}{\rho_L} \left| p_L - p_Z \right|} \ \Big] \tag{7}$$

where: $A_z^{(i)}$ – variable surface area of fuel outflow through the outlet aperture,

V_L – volume of the inlet chamber,

p_k – pressure in the combustion chamber (back-pressure).

$$\varepsilon_z^{(i)} = \begin{cases} 1 & \text{where } t \in \left\langle t_A^{(i)}, \ t_B^{(i)} \right\rangle \ \ i \ \ p_z > p_k \\ 0 & \text{in all other cases} \end{cases}$$

$t_A^{(i)}, t_B^{(i)}$ – aperture opening times,

other symbols as previously defined in parts 2.1 and 2.4.

Equation (7) is a non-linear first-order ordinary differential equation.

Modelling of hydrodynamic phenomena in the injector presents a number of difficulties. An essential aspect for a correct model of the rail is the definition of a rule describing the flow of fuel from the rail to the combustion chamber. In the algorithm presented here, it was

assumed that further apertures are opened at intervals equal to 180° of pump shaft rotation. It is also important to choose the values for a number of coefficients necessary to conduct a proper quantitative assessment of the phenomena taking place. Here, coefficients of hydraulic resistance, flow rate coefficients and coefficients of resistance to movement of moving parts can all be mentioned. The values of those quantities vary in response to fuel pressure, which complicates their determination. Furthermore, when model ling electronically-controlled injection systems, solenoid control valves must be considered, which requires familiarity with further quantities, especially the properties of the materials used. The values of certain quantities are sometimes difficult to estimate, and thus it was decided to develop and empirical model for the flow of fuel from the atomizer, based on function $A_z^{(i)}$. Values of characteristic times obtained from the analysis of injection events were used.

A basic observation made during the experiments was that the real trace of the needle lift, as well as function $A_z^{(i)}$ deviate from theory, in which the following are defined:

$t_B^{(i)}$ – given opening time, $t_0^{(i)}$ – given pause time.

Most importantly, it can be stated that the real opening time $\overline{t_B}$ is greater than the given $t_B^{(i)}$ by an approximately constant magnitude, denoted as $t_d^{(i)}$ – injection delay time. The delay time considers differences between the given and realized injection time; the value was determined experimentally. The trace of function $A_z^{(i)}$ also took forms more closely approximation to a parabola than to the theoretical profile in the form of a rectangular function.

Depending on the value of $t_0^{(i)}$ and $t_d^{(i)}$, the following two cases can be obtained:

$$t_d^{(i)} < t_0^{(i)} \qquad \text{and} \qquad t_d^{(i)} > t_0^{(i)}$$

taken into consideration in the computer programme developed in this work.

In the programme calculating the injection process, the possibility of entering a pressure value below the injection should not commence was included (this is equivalent to the pressure opening the injector). This is protection against calculating injection parameters in cases where the quality of the atomisation process (not analysed via this model) could prove to be unsatisfactory.

3. Numerical integration of the differential equation system

The majority of the methods for integration of systems of ordinary differential equations require the insertion of higher-order equations to first-order equations. Thus, equations (3) and (4) were inserted into the appropriate two first-order equations. Equations (1), (3), (4) and (7) were then saved in the form of a system of first-order equations, of form:

$$\dot{X} = F(t, X) \tag{8}$$

where F is a vector function and X:

$$X = \left[p_L, p_p^{(1)}, ..., p_p^{(3)}, h_g^{(1)}, \dot{h}_g^{(1)}, ..., h_g^{(3)}, \dot{h}_g^{(3)}, h_k^{(1)}, \dot{h}_k^{(1)}, ..., h_k^{(3)}, \dot{h}_k^{(3)}, p_z \right]^T$$

is a vector with m = 17 elements.

It is therefore necessary to integrate the system m = 17 first-order ordinary differential equations. To that end, the Runge-Kutte 4th order method was employed, using constant integration steps.

4. Initial conditions, further points

Calculations were conducted assuming that for the first instant (t = 0) all pressures are equal to the pressure of the inflowing (feed) fuel p_d and all lift values and speeds are zero, i.e.:

$$X_{t=0} = \left[p_d, p_d, p_d, p_d, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, p_d \right]^T \tag{9}$$

It was also assumed that the first section of the high-pressure pump is the first to start working, the other two sections commencing after (respectively) 120 and 240 degrees of pump shaft rotation. It was therefore assumed that angles $\alpha^{(i)}$ are described by the dependency:

$$\alpha^{(i)} = \begin{cases} 0 & \text{where } t < t_i \\ \omega(t - t_i) & \text{where } t > t_i \end{cases} \tag{10}$$

where: ω - angular velocity,

$$t_i = 0 \quad \text{dla } i = 1, \qquad t_i = \frac{2\pi}{3\omega} \quad \text{dla } i = 2, \qquad t_i = \frac{4\pi}{3\omega} \quad \text{dla } i = 3.$$

As empirical research was conducted for the assumed operating conditions for the analysis of the fuel supply system, results obtained from the computer simulation can be considered only after multiple work cycles (φ > 720°), since during the initial phase the computation is too sensitive to the impact of initial conditions (9) and shifts t_i in equation (10).

As previously mentioned, the outlet apertures on the rail are further activated (cyclically) every:

$$\Delta T = t_B^{(i)} - t_A^{(i)} = \frac{\pi}{\omega} \tag{11}$$

The programme for modelling operation of a common rail injection system developed based on the dependencies presented above, facilitates the calculation of the pressure trace in the pump chambers, the inflow chamber and the fuel rail, as well as the lift of pistons and movable valve elements. The complete fuel dose in the injection and the fuel flow rate through particular injection apertures are computed. Injection event traces can be designated for both non-split and split doses and for varying pause time values.

Verification calculations were conducted for a fuel supply system with a cylindrical high pressure rail (rail). Comparisons were carried out for the following: split injection – pilot event 450 µs, pause 600 µs, main event 450 µs, at a set rail pressure of 700 bar, pump rotational speed 695 rpm, injection order 1 – 2 – 3 – 4. Differences between the calculated and measured values ranged from 2.4% to 7.7%, depending on the injector selection group. Such differences mainly result from the simplifications adopted in the model, as high pressure hoses and injector assemblies are not included. The dose delay time and its dependence on fuel pressure have a significant impact on the dose.

5. Impact of rail geometric parameters on the injection process

Using the considered injection system model, using the program calculating the injection event, calculations were performer for various setpoints of the injector control signal. A split (two-part) fuel dos-

ing strategy was considered. The calculations aimed to provide qualitative and quantitative evaluations of the influence of test quantities on the injection parameters. The considerations regarding the influence of geometrical parameters of the fuel container on the injection process are presented here.

The rail is a relatively simple element in terms of its design, yet it plays a key role in the limitation of the propagation of pressure waves. A correctly chosen capacity ensures continuity of dosing during abrupt changes in engine operating parameters. As previously mentioned, a cylindrical high pressure rail used on the injection system of a compression-ignition engine (of swept volume 1700 cm^3) was adopted for carrying out model calculations.

Using the model, reviews of the influence of the length, diameter and volume of the rail on the fuel supply parameters were conducted. Calculations were conducted: for fixed rail diameter and variable rail length and fixed rail length and variable rail diameter. Changes to the trace of the injection process, the fuel dose, the angle of the start of injection and the angle of injection duration were analysed.

Evaluation of the impact of rail length

Figures 2 and 3 present the traces of injections from an injector calculated for constant diameter and various rail lengths. The dashed purple line presents results for the basic rail length used by the engine manufacturer (201.4 mm). Here, the injection angle is 11° and this remains the same for all cases. However, the angle of start of injection changes; for the considered range of lengths the range of changes amounts to 8° of pump shaft rotation. This is a relatively important change in an important injection parameter, which must be considered in the design of algorithms controlling engine operation. Such changes result – above all – from the means of controlling the injector in the model, which enables it to open at a given pressure value.

Together with the increase in rail length, the mean value of rail pressure barely changes (0.02%) and such changes are practically unnoticeable. Similarly insignificant changes result from extreme flow rates of fuel from the atomizer.

Conversely, there are differences between the maximum and minimum pressure values. For a rail of length 160 mm, the difference amounts to 77 bar, for a rail of length 201.4 mm 62.6 bar, and for 250 mm as little as 51.9 bar. Such changes affect the behaviour of fuel in the rail.

The changes in injection process parameters presented here result – above all – from increases in the volume of the element under consideration. As the calculations were made with control system settings unchanged, increases in the volume chase progressively later

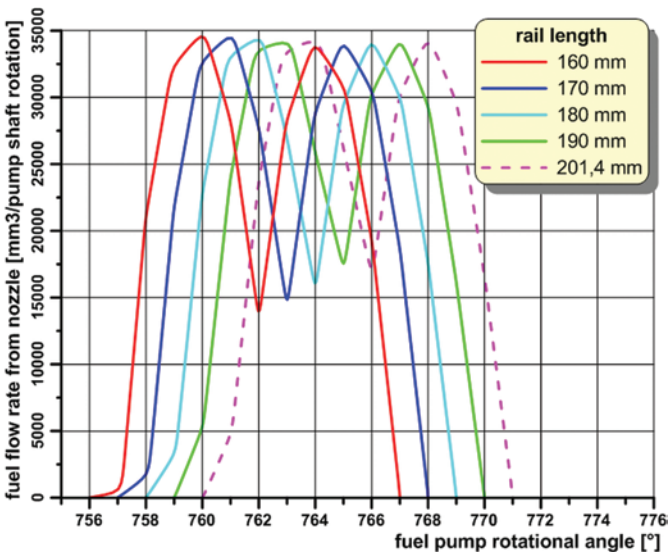


Fig. 2. Calculated injection traces for rail lengths 160 mm – 201.4 mm

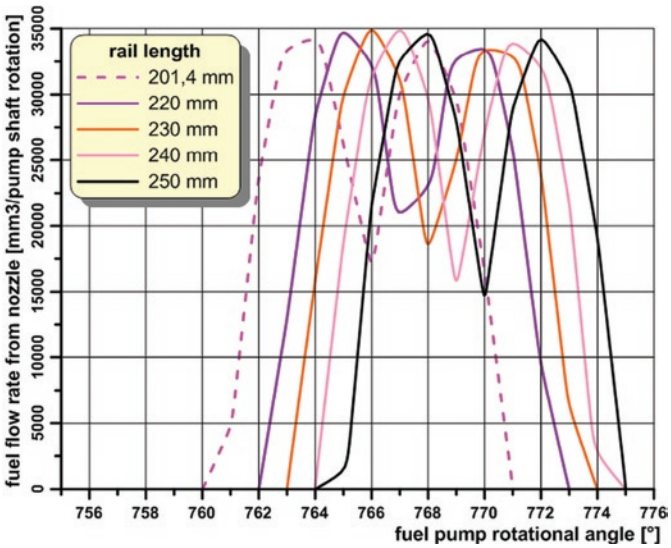


Fig. 3. Calculated injection traces for rail lengths 201.4 mm – 250 mm

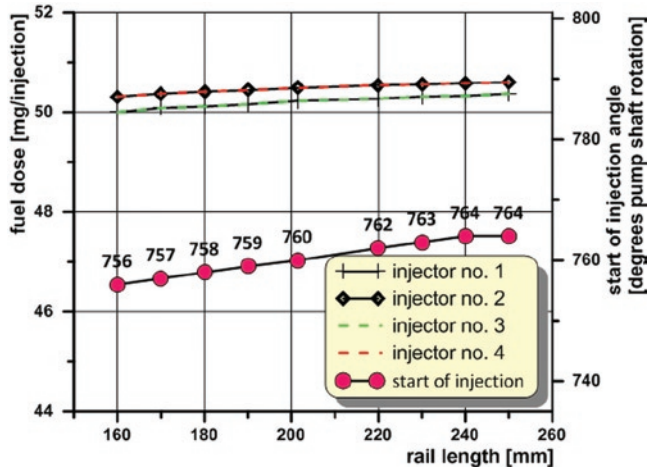


Fig. 4. Calculated fuel doses and angles of start of injection for various rail lengths

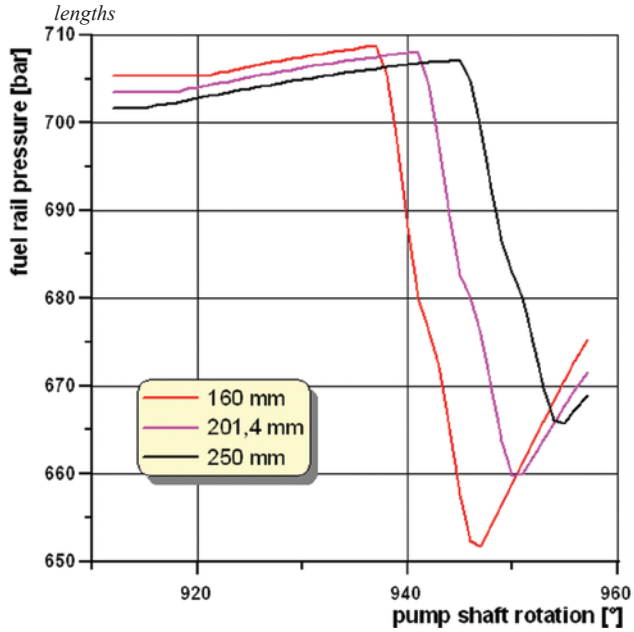


Fig. 5. Calculated pressure changes in the rail during the injection event, together with angles of start of injection for selected rail lengths

achievement of the required pressure level. This is the delay of start of injection (Fig. 4). As the injection duration time does not change,

the end of the injection event also occurs later and later, which leads to ever smaller differences in the pressure between the start and end of injection (Fig. 5). Thus, a small (0.7%) increase in the fuel dose was noted.

Evaluations of the influence of rail diameter

The changes in the high-pressure rail length analysed above had a linear character and changed the injection parameters in the same way. The situation is somewhat different where injections are considered for constant rail length and variable rail diameter. In such considerations the direction of change is similar, yet changes in the volume are significant and occur non-linearly, in proportion to the given diameter raised to the second power. In order to obtain a full picture of changes in injection parameters, a wide range of diameter values was used, from the smallest (corresponding to the diameter of the injection hose) up to 20 mm – thus being values greater than those used in the majority of fuel rails used in passenger car engines. With regard to changes in rail diameter, it can be seen that a greater proportion of the energy delivered to the rail is used in the process of compressing the fluid. The greater quantity of fuel in the rail absorbing the delivered energy causes significant delays in the start of injection – by as much as 64° (Fig. 6, also Fig. 8), at a constant injection angle value of 11°. However, in this case an important role is played by the control of

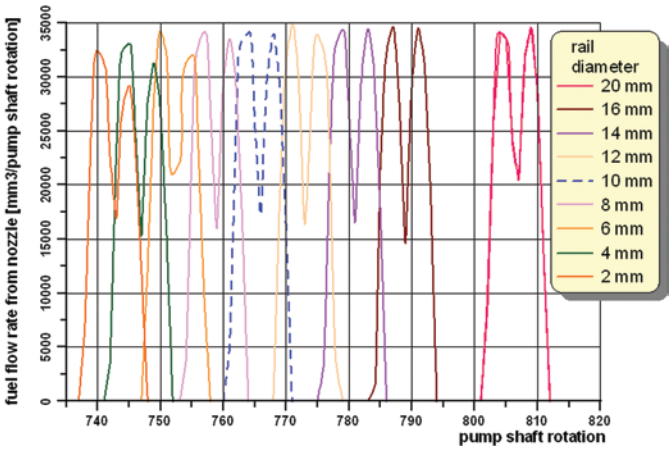


Fig. 6. Calculated injection traces for various rail diameters

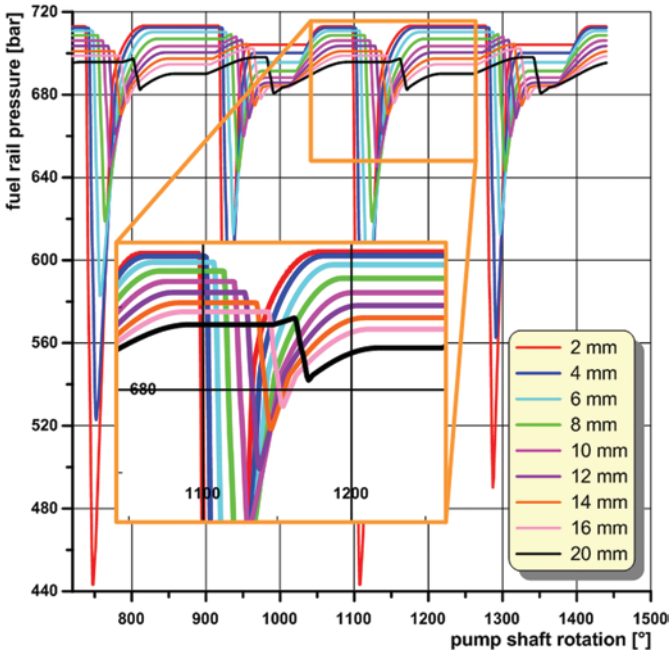


Fig. 7. Calculated pressure traces for various rail diameters

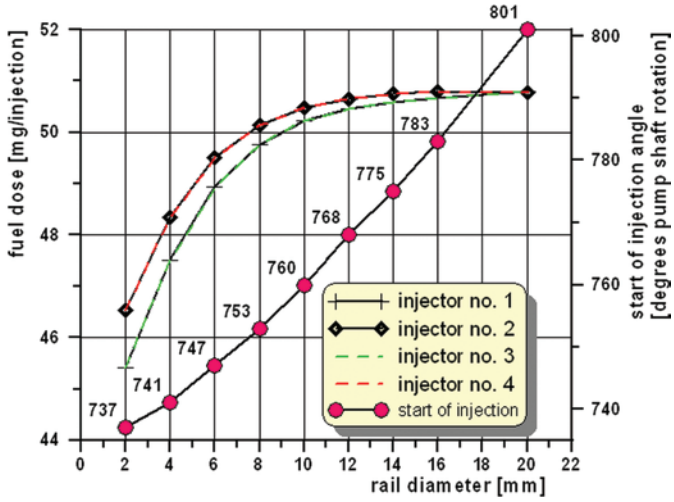


Fig. 8. Calculated dose and angle of start of injection values for various rail diameters

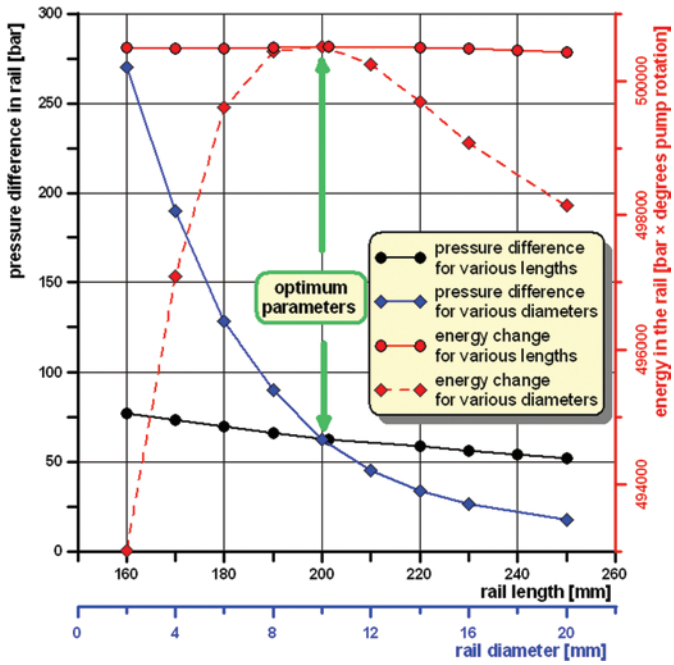


Fig. 9. Calculated pressure differences and energy changes for various rail dimensions

injection opening. Mean rail pressure values change by some 1.5% and do not reflect changes in the rail during the injection process, especially in the case of the smallest diameters, where large pressure changes occur (Fig. 7). This causes greater changes in the fuel dose and angle of start of injection. With an increase in rail diameter, significant differentiation in the flow rate of fuel from the atomizer is in fact not observed (Fig. 6). Such changes can be observed only for the smallest rail diameters – i.e. where the greatest pressure drops occur. The shape of the injection event does not change. Pressure differences (Fig. 7) have an impact on the fuel dose delivered.

For large pressure drops, part of the process occurs at low pressure values, leading to a lower quantity of fuel (Fig. 8). Over the entire range considered the change in the dose was significant, as it was 11.8%.

The modelling analysis results presented above do not fully resolve all questions. Additional calculations were made, the results of which were taken into consideration in a qualitative analysis of the influence of the aforementioned fuel rail operating parameters on the injection process (Table 1).

Table 1. Qualitative evaluations of investigated parameters

Parameter	Value range	Impact of test parameter on:		
		fuel dose [mg]	injection duration angle [°]	angle of start of injection [°]
Pause time	200 µs-900 µs	+	+ + +	- -
Injection delay time	100 µs-900 µs	+ + +	+ + +	- -
Rail length (with spigot spacing changed proportionally)	160 mm-250 mm	+ -	- -	+
Rail diameter	2 mm-20 mm	+ + +	- -	+ + +

Legend: *+ + + high impact*
+ significant impact
+ - insignificant impact
- - no impact

The aforementioned differences between the maximum and minimum rail pressure values are presented in Fig. 9. Differences for various lengths are shown in black and compared with differences occurring as a result of changes in rail length (blue line). It can be seen that as the length and diameter of the rail increase, the pressure differences resulting from the injection process are reduced, and the impact of changes in the rail diameter is significantly greater.

If it may be assumed that a measure of the available energy in the fuel before the injection is the area under the rail pressure trace curve, then that quantity changes for various rail lengths to a degree equal to changes in the mean pressure value – that is to say, very little (Fig. 9, solid red line). The same figure shows changes in the energy of the fuel in the rail (red line). While changes in rail length (within the considered range) do not cause significant changes in this quantity, the situation is completely different for changes in diameter (Fig. 9, dashed red line). The trace reaches its maximum for a rail of diameter 10 mm and length 201.4 mm. For such a dimensional configuration the fuel energy before the injection is the greatest and can be property used to prepare the fuel-air mixture. The values given here as optimal were taken by the manufacturer of the fuel system under analysis and employed in engines of swept volume 1700 cm^3.

It should, however, be underlined that the rail was modelled in a simplified manner, without considering wave phenomena. A real fuel rail is subject to the laws of wave propagation; local areas of pressure higher or lower than the given value occur, which can influence the dosing process to a great degree. Following consideration of such effects, changes in the quantitative evaluations of the dependencies presented here may occur.

6. Summary

Modelling plays a significant role in the design and selection of machine parts. It allows the implementation time of the designed system to be significantly shortened, as well as to adapt it to the building stage in parallel with the design process. In order for the mathematical model to best reflect real phenomena, a correct physical model of the test system should be prepared. It is obvious that the degree of simplification of the modelled system will affect the accuracy of the results of calculations, but in many cases the use of simplifications is necessary, due to the complexity of the mathematical model, which would increase the required calculation time and thus reduce the efficiency of the program.

The developed injection process model, applicable to the commonly used fuel injection system in common rail compression-ignition engines, allowed determination of the relationships between the parameters of the test system. From the obtained test results, it is possible to distinguish the factors which have the greatest impact on the fuel dose, the injection trace, the start angle and the duration of injection. The impacts of the analyzed quantities on injection parameters varied; they can be summarized as follows.

- The pressure trace has a significant impact on the en tire injection process and fuel dose. Increased rail pressure causes changes in the flow rate of fuel leasing the atomizer, which translates into an increase in fuel expenditure.
- For a given length within the range of considered values, the diameter of the rail Has a significant influence on the angle of the start of injection. Increases in this value cause increases in the start angle and decreases in the quantity of fuel dosed. This results from the volume and compressibility of the fuel, since greater volume cause extended reaction times to the signal controlling the rail pressure.
- Changes in rail length for constant rail diameter have an insignificant influence on the angle of the start of injection. Greater rail length increases the angle of the start of injection, but to a lesser degree than changes in diameter. This results from the lesser increase in fuel volume. This quantity does not affect the injection duration angle.
- Changes in rail diameter and length at constant rail volume do not have Any influence on the injection parameters analysed here.

Taking into account the simulation results obtained, as well as differences between the results of calculations and the values measured on the test bench, it has been confirmed that the values are very similar. However, as always, there are discrepancies between the real system and the model. The values of these discrepancies permit assessment of the quality of the model and its susceptibility to changes in the set values. Despite the considerable complexity of the algorithms and the large number of parameters that can be changed, the model is relatively predictable in terms of the results it generates. This feature allows for quick model tests and selection of suitable initial parameters that will allow the desired trace of the injection process and fuel dosage to be obtained. The computer calculation program, developed on the basis of the physical model presented in this paper can be qualified as reflecting the studied parameters of the injection system well. Due to some simplifications adopted in the mathematical model, there are differences in the results of calculations and results from measurements, but they do not significantly change the results. The evident imperfection is the limited split of the fuel dose into parts.

In the current version of the program, only two parts can be made. In the trace of further work, the program and model should be adapted to current requirements and, using experimental results, empirical and computational models should be developed, taking into account both the possibility of dividing the injection into a larger number of parts and including a larger number of control parameters.

References

1. Ahlin K. Modelling of pressure waves in the common rail Diesel injection system. Thesis performed in Automotive Systems at Linköping Institute of Technology, LiTH-ISY-EX-3081, 2000.
2. Amoia V, Ficarella A, Laforgia D, De Matthaeis S, Genco C. A theoretical code to simulate the behavior of an electro-injector for Diesel engines and parametric analysis. 1997, SAE Technical Paper 970349, https://doi.org/10.4271/970349.
3. Arcoumanis C, Baniasad M. S. Analysis of consecutive fuel injection rate signals obtained by the Zeuch and Bosch Methods. 1993, SAE Technical Paper 930921, https://doi.org/10.4271/930921.
4. Gautier C, Sename O, Dugard L, Meissonnier G. An LFT Approach to H_ Control Design for Diesel Engine Common Rail Injection System. Oil & Gas Science and Technology, Rev. IFP, 2007,62(4): 513-522, https://doi.org/10.2516/ogst:2007046.
5. Gautier C, Sename O, Dugard L, Meissonnier G. Modelling of a Diesel Engine Common Rail Injection System. IFAC 16th World Congress, Prague, 2005, https://doi.org/10.3182/20050703-6-CZ-1902.01919.
6. Huhtala K, Vilenius M. Study of a common rail fuel injection system. 2001, SAE Technical Paper 2001-01-3184, https://doi.org/10.4271/2001-01-3184.
7. Lino P, Maione B, Pizzo A. Nonlinear modelling and control of a common rail injection system for diesel engines. Applied Mathematical Modelling, 2007, 31(9): 1770-1784, https://doi.org/10.1016/j.apm.2006.06.001.
8. Ricco M, De Matthaeis S, Olabi A.G. Simulation of the magnetic properties for common rail electro-injector. Journal of Materials Processing Technology, 2004, 155(1): 1611-1615, https://doi.org/10.1016/j.jmatprotec.2004.04.343.
9. Schuckert M, Schultze L, Tschöke H. Zur Auslegung von Common-Rail Diesel-Einspritzsystemen. MTZ, 1998, 59(12): 800-806, https://doi.org/10.1007/BF03226481.
10. Seykens X.L.J, Somers L.M.T, Baer R.S.G. Detailed Modelling Of Common Rail Fuel Injection Process. Journal of Middle European Construction and Design of Cars, 2005, 3(2-3): 30-39.
11. Shinohara Y, Takeuchi K, Hermann O. E, Lumen H. J. Common-Einspritzsystem mit 3000 bar. MTZ, 2011, 72(1): 10-15, https://doi.org/10.1365/s35146-011-0005-7.

Tomasz KNEFEL
Jacek NOWAKOWSKI
Department of Combustion Engines and Vehicles
University of Bielsko-Biala
ul. Willowa 2, 43-309 Bielsko-Biała, Poland

E-mails: knefel@ath.bielsko.pl, jnowakow@ath.bielsko.pl

Article citation info:
SELECH J, ANDRZEJCZAK K. An aggregate criterion for selecting a distribution for times to failure of components of rail vehicles. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 102–111, http://dx.doi.org/10.17531/ein.2020.1.12.

Jarosław SELECH
Karol ANDRZEJCZAK

AN AGGREGATE CRITERION FOR SELECTING A DISTRIBUTION FOR TIMES TO FAILURE OF COMPONENTS OF RAIL VEHICLES

ZAGREGOWANE KRYTERIUM WYBORU ROZKŁADU CZASU DO USZKODZENIA ELEMENTÓW POJAZDÓW SZYNOWYCH*

This paper presents an aggregate method of selecting a theoretical cumulative distribution function (CDF) for an empirical CDF. The method was intended to identify the time of reliable operation of a renewable technical object by applying three criteria based on the following statistics: the modified Kolmogorov–Smirnov (MK-S) statistic, the mean absolute deviation of the theoretical CDF from the empirical CDF, and a statistic calculated on the basis of a log-likelihood function. The values of these statistics were used to rank eleven probability distributions. The data for which calculations were made concerned failures of the driver's cab lock recorded during five years of operation of a fleet of 45 trams. Before calculating the statistics, the empirical CDF of the examined component was determined using the Kaplan–Meier estimator, and then, using the method of Maximum Likelihood Estimation, the parameters of the analysed theoretical distributions were estimated. The theoretical distributions were then ranked according to the values obtained for each of the assumed criteria: the lower the value for a given criterion, the higher the ranking position, indicating a better fit according to that criterion. Then, based on the three rankings and on weights assigned to the individual criteria, an aggregate criterion (referred to as DESV) was implemented to select the best-fitting probability distribution. The method assumes that the lowest DESV value corresponds to the best-fitting theoretical distribution. In the case of the examined component, this was found to be the generalised gamma distribution. It is shown that if the final decision is based on the aggregate criterion, which takes into account the three criteria for goodness of fit, the reliability of the estimation of the time-to-failure distribution increases, and thus mistakes resulting from the use of only one of the criteria can be avoided.

***Keywords**: time to failure, estimation of probability distribution, reliability of rail vehicles.*

W pracy przedstawiono zagregowaną metodę doboru dystrybuant hipotetycznych do dystrybuanty empirycznej. Metoda miała na celu identyfikację czasu niezawodnej pracy odnawialnego obiektu technicznego poprzez zastosowanie trzech kryteriów, w których użyto następujących statystyk: zmodyfikowanej statystyki Kołmogorowa-Smirnowa (MK-S), statystyki średniego odchylenia bezwzględnego dystrybuanty hipotetycznej od empirycznej oraz statystyki obliczanej na podstawie zlogarytmowanej funkcji wiarygodności. Wartości tych statystyk posłużyły do rangowania jedenastu rozkładów prawdopodobieństwa. Dane dla których dokonano obliczeń dotyczyły uszkodzeń zamka kabiny motorniczego jakie odnotowano w ciągu pięciu lat użytkowania floty 45 tramwajów. Przed obliczeniem statystyk wyznaczono dystrybuantę empiryczną badanego elementu przy pomocy estymatora Kaplana-Meiera, a następnie przy użyciu metody największej wiarygodności oszacowano parametry uwzględnionych w badaniach rozkładów hipotetycznych. Po wyznaczaniu parametrów nastąpiło rangowanie rozkładów hipotetycznych według wartości otrzymanych dla każdego z przyjętych kryteriów, im mniejsza wartość dla danego kryterium tym wyższa pozycja w rankingu, świadcząca o lepszej jakości dopasowania według danego kryterium. Po ustaleniu rankingu według kryteriów zgodności, każdemu z kryteriów zgodności dopasowania dystrybuant modelowych do empirycznej nadano wagi. Następnie na podstawie uzyskanych trzech rankingów oraz wag nadanych poszczególnym kryteriom zgodności wyznaczana jest zagregowana miara zgodności (oznaczona DESV), która służy do wyznaczania najlepszego rozkładu prawdopodobieństwa. W prezentowanej metodzie przyjęto, że najmniejsza wartość DESV wyznacza najlepiej dopasowany rozkład hipotetyczny. W przypadku badanego elementu rozkładem tym okazał się uogólniony rozkład gamma. Pokazano, że na podstawie zagregowanego kryterium uwzględniającego trzy statystyki zgodności dopasowania zwiększa się wiarygodność estymacji rozkładu czasu pracy do uszkodzenia, unikając tym samym błędów jakie można popełnić uzależniając się tylko od jednej z nich.

***Słowa kluczowe**: czas do uszkodzenia, estymacja rozkładu prawdopodobieństwa, niezawodność pojazdów szynowych.*

1. Introduction

In traditional methods of estimating the parameters of the time-to-failure distribution of a technical object or its components, a specific distribution class is assumed *a priori*. The purpose of this article is to present the results of a procedure to identify the best-fitting probability distribution model for the time to failure of a renewable technical object using an aggregate criterion. The research concerns components of currently operated rail vehicles of a uniform type that belong to a fleet maintained by the operator. Empirical data obtained during the operation of the vehicles are incomplete, since the vehicles were operational at the end of the data acquisition period. Thus, the authors did not have complete data on the times to failure of all components of the analysed vehicles. Therefore, it was necessary to use statistical methods taking account of censored data. Given a suitably prepared database of repairs to vehicles in the fleet, it is relatively easy to de-

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

Table 1. Density functions and parameters to be estimated

Distribution model	Probability density function	Distribution parameters
Exponential	$f(t;\lambda)=\lambda e^{-\lambda t},\ t\geq 0,\lambda>0$	$\frac{1}{\lambda}$ – scale parameter
Two-parameter exponential	$f(t;\lambda,\gamma)=\lambda e^{-\lambda(t-\gamma)},\ t\geq\gamma,\lambda>0$	$\frac{1}{\lambda}$ – scale parameter γ – location parameter
Normal	$f(t;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2},\ t\in\mathbb{R},\mu\in\mathbb{R},\sigma>0$	μ – expected value σ – standard deviation
Lognormal	$f(t;\mu',\sigma')=\frac{1}{t\cdot\sigma'\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\ln(t)-\mu'}{\sigma'}\right)^2},\ t>0,\mu'\in\mathbb{R},\sigma'>0$	μ' – expected value of normally distributed ln T σ' – standard deviation of ln T
Two-parameter Weibull	$f(t;\beta,\eta)=\frac{\beta}{\eta}\cdot\left(\frac{t}{\eta}\right)^{\beta-1}\cdot e^{-\left(\frac{t}{\eta}\right)^{\beta}},\ t\geq 0,\beta>0,\eta>0$	η – scale parameter β – shape parameter
Three-parameter Weibull	$f(t;\beta,\eta,\gamma)=\frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1}e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}},\ t\geq\gamma,\beta>0,\eta>0,\gamma\in\mathbb{R}$	η – scale parameter β – shape parameter γ – location parameter
Gamma	$f(t;\mu,\kappa)=\frac{\exp\left(\kappa(\ln(t)-\mu)-\exp\left(\ln(t)-\mu\right)\right)}{t\Gamma(\kappa)};\ t>0,\mu\in\mathbb{R},\kappa>0$	e^{μ} – scale parameter κ – shape parameter
Generalised gamma	$f(t;\theta,\beta,\kappa)=\frac{\beta}{\Gamma(\kappa)\cdot\theta}\cdot\left(\frac{t}{\theta}\right)^{\kappa\beta-1}\cdot e^{-\left(\frac{t}{\theta}\right)^{\beta}},\ \theta>0,\beta>0,\kappa>0$ Reparameterisation: $\mu=\ln(\theta)+\frac{1}{\beta}\ln\left(\frac{1}{2}\right);\ \sigma=\frac{1}{\beta\sqrt{\kappa}};\ =\frac{1}{\sqrt{\kappa}}$ $f(t;\mu,\sigma,\lambda)=\begin{cases}\frac{\lvert\lambda\rvert}{\sigma\cdot t}\cdot\frac{1}{\Gamma\left(\frac{1}{2}\right)}\cdot\exp\left[\frac{\lambda\frac{\ln(t)-\mu}{\sigma}+\ln\left(\frac{1}{\lambda^2}\right)-\exp\left(\frac{\ln(t)-\mu}{\sigma}\right)}{\lambda^2}\right] & if\ \lambda>0\\ \frac{1}{t\cdot\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(t)-\mu}{\sigma}\right)^2\right) & if\ \lambda=0\end{cases}$ $t\geq 0,\mu\in\mathbb{R},\lambda\geq 0,\sigma>0,$	θ – scale parameter β – shape parameter κ – shape parameter
Logistic	$f(t;\mu,\sigma)=\frac{\exp\left(\frac{t-\mu}{\sigma}\right)}{\sigma\left(1+\exp\left(\frac{t-\mu}{\sigma}\right)\right)^2},t\in\mathbb{R},\ \mu\in\mathbb{R},\sigma>0$	σ – scale parameter μ – location parameter
Loglogistic	$f(t;\mu,\sigma)=\frac{\exp\left(\frac{\ln(t)-\mu}{\sigma}\right)}{\sigma t\left(1+\exp\left(\frac{\ln(t)-\mu}{\sigma}\right)\right)^2},\ t>0,\ \mu\in\mathbb{R},\sigma>0$	μ – scale parameter σ – shape parameter
Gumbel	$f(t;\mu,\sigma)=\frac{1}{\sigma}\exp\left(\left(\frac{t-\mu}{\sigma}\right)-\exp\left(\frac{t-\mu}{\sigma}\right)\right),\ \sigma>0$	μ – location parameter σ – scale parameter

termine basic reliability characteristics of failed components [25]. However, the selection of a good criterion for the fit of a distribution of times to failure of the components becomes an issue. This problem is the subject of the research presented in this paper, which concerns the use of an aggregate criterion for determining the best-fitting time-to-failure distributions for selected components of a rail vehicle [33]. The research results are presented in the form of a ranking of the fit of selected families of distributions based on the aggregate criterion.

In the study of technical objects, different probability distribution families are used as models of time to failure [17]. The most commonly used distributions in Life Data Analysis (LDA) are the normal, exponential and Weibull distributions [19, 10]. In this study, apart from the aforementioned distributions, the authors also verified the possibility of using other, less common distributions, whose goodness of fit to the empirical data proved superior in many cases to the more common distributions. These are the lognormal, gamma, generalised gamma, logistic, loglogistic and Gumbel distributions [22]. The density functions of these distributions and their parameters are listed in Table 1. In the case of the generalised gamma distribution, for easier parameter estimation, the density function is also given in reparameterised form [20].

The parameters of these distributions can be estimated using analytical, numerical and graphical methods [16, 26, 29]. The most commonly used methods include the method of moments, Maximum Likelihood Estimation, the least squares method, the method of probability plotting, and the probability plot correlation coefficient (PPCC) method [1, 38, 32]. In engineering practice, the most commonly used are numerical and graphic methods executed with specialised IT tools [12, 39]. Based on the operational data and selection of the estimation method, parameters (shape, scale, location) are estimated for selected families of probability distributions [28, 37]. Having estimated various distributions, it is possible to indicate which of them is the best fitted to the empirical data in the sense of the lowest sum of squares of deviations.

The proposed methodology for identifying the time to failure of a selected vehicle component uses all available data on times (mileages) between failures of the component in all vehicles of the analysed fleet. This includes the case where the component was operational at the time when data acquisition was ended; the time to failure of such a component is said to be right-censored. A method of preparing statistical data based on the operational database has been developed in the articles [3, 2].

Instead of the traditional single-criterion selection of the best-fitting family of probability distributions, the authors propose to use an aggregate criterion that includes three measures of the fit of theoretical distributions. This criterion takes into account a ranking of the fit of individual probabilistic models to the empirical data, including right-censored operational data for the vehicle fleet.

In the aggregate method, the choice of a distribution is preceded by a ranking of distributions for three goodness-of-fit criteria. The parameters of selected distribution families were estimated using the Weibull++ Distribution Wizard module, which – after performing the appropriate calculations – ranks the distributions starting from that with the highest goodness of fit. However, before the fit of the distributions is examined, the CDF or reliability function of the empirical distribution is determined by the Kaplan–Meier method, and then the parameters of the theoretical distributions are determined by the method of Maximum Likelihood Estimation (MLE).

The next step is to determine the statistics of the goodness of fit of the theoretical CDFs to the empirical CDF, denoted F_n. On this basis, a ranking is made of eleven distributions, listed in Table 1, that are used in the survival analysis [18, 30]. Provided that the assumptions are met, the rankings of distributions by goodness of fit are compiled independently according to three criteria, using the modified Kolmogorov–Smirnov (MK-S) statistic, the statistic of the mean absolute deviation of the theoretical CDF from the empirical CDF, and the value of the log-likelihood function [23].

The final ranking of distributions is based on the rankings obtained using these three criteria, taking into account the weights assigned to each of them. After assigning weights to the criteria, the final Distribution Estimation Values (DESV) are calculated, indicating the best-fitting distribution according to the aggregate criterion. The scheme of successive calculation steps in the aggregate method of ranking distributions is shown in Fig. 1.

According to this scheme, in the first step, based on the obtained data and analysing the length of the observation time (right-censored), the survival function parameters were estimated with the Kaplan–Meier estimator and an empirical CDF was determined [7]. Then, to determine the parameters of the eleven theoretical distributions listed in Table 1, the method of Maximum Likelihood Estimation was used [15, 11].

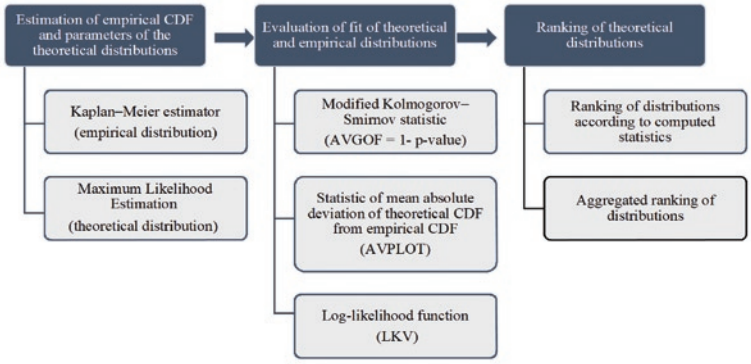


Fig. 1. Aggregate criterion for ranking distributions

In the second step, for each of the eleven distributions, the goodness-of-fit statistics are used to test the null hypothesis:

$$H_0 : T \sim F \tag{1}$$

stating that the time to failure T of the analysed vehicle component has a probability distribution with the CDF F with the estimated parameters. This evaluation is based on a random sample $T_1, T_2, \ldots, T_n$ concerning times to failure of the component. In this paper, the times to failure of the examined component are expressed in terms of kilometres travelled, as in the paper [2].

2. Criteria for ranking theoretical distributions

Among the applied goodness-of-fit criteria, a particular role is played by the modified Kolmogorov–Smirnov statistic (AVGOF, average goodness of fit), which evaluates the statistical difference between the values of the empirical and theoretical CDFs. The particular role of this statistic results from the fact that it is highly sensitive to local deviations. In addition, it can be used even with a small amount of data and with unknown parameters of the theoretical distribution. The use of the MK-S statistic is therefore necessary when the parameters of the tested distributions need to be estimated.

Because the distributions of MK-S statistics depend on a theoretical distribution family whose parameters are estimated, a critical value, at which the null hypothesis is rejected, is determined for each distribution [30]. Analytical determination of the critical value is often difficult or even impossible, and hence it is obtained using the Monte Carlo method [6, 18].

The MK-S statistic used to test the fit of a theoretical distribution to the empirical distribution uses the statistic D_{max}, defined as the maximum of the absolute difference between the value of the empirical CDF $F_n(t)$ and a matched theoretical CDF $F(t)$, and given by the formula [18]:

$$D_{max} = \max_{1 \le i \le n} \left| F_n(t_i) - F(t_i) \right| \tag{2}$$

where:

D_{max} is the value of the statistic;
n is the sample size;
$F_n(t_i)$ is the value of the empirical CDF;
$F(t_i)$ is the value of the theoretical CDF.

The critical value D_{CRIT} in the modified Kolmogorov–Smirnov statistic is determined by the Monte Carlo method, as already mentioned, due to the difficulty of the calculations.

The MK-S statistic is used to determine the probability of rejection of the null hypothesis, i.e. the probability of the event $D_{CRIT} < D_{max}$. Hence, in the case of the first criterion, the basis for ordering theoretical distributions is the probability:

$$\mathrm{P}\left(D_{CRIT} < D_{max}\right) \tag{3}$$

The higher the value of the statistic D_{max}, the more significant is the difference between the theoretical distribution defined by the CDF F and the empirical distribution with the CDF F_n. Because the critical value D_{CRIT} is determined by the Monte Carlo method through m-tuple generation of n time-to-failure values $t_{s1}, t_{s2}, \ldots t_{sn}$, for which simulation CDFs $F_s(t_{si}), s = 1,2,\ldots,m$ are created, and maximum differences with the values of the theoretical CDF are determined for each of these:

$$d_{max,s} = \max_{1 \le s \le m} \left| F_s(t_{si}) - F(t_{si}) \right|, s = 1,2,\ldots,m \tag{4}$$

the critical value D_{CRIT} is estimated as the arithmetic mean d_{CRIT} defined by the formula [6]:

$$\widehat{D_{CRIT}} = d_{CRIT} = \frac{1}{m} \sum_{s=1}^{m} d_{max,s} \tag{5}$$

Finally, in the MK-S criterion for the goodness of fit of distributions we assume:

$$\mathrm{AVGOF}\left(F_n, F\right) = 100 \cdot \mathrm{P}\left(d_{CRIT} < D_{max}\right) \tag{6}$$

Large values of AVGOF, close to 100, indicate that there is a significant difference between the theoretical distribution and the empirical data. Hence, the lower the value of the statistic AVGOF, the better the fit of the theoretical distribution.

In the case of the second goodness-of-fit criterion, the mean absolute deviation of the theoretical CDF from the empirical CDF is examined, and the statistic used to assess goodness of fit, denoted AVPLOT (average plot fit), is determined according to the formula:

$$\mathrm{AVPLOT}\left(F_n, F\right) = 100 \frac{1}{n} \sum_{i=1}^{n} \left| F_n(t_i) - F(t_i) \right| \tag{7}$$

where:

n is sample size;
$F_n(t_i)$ are values of the empirical CDF;
$F(t_i)$ are values of the theoretical CDF.

This criterion, unlike the MK-S criterion, is not sensitive to local deviations, but takes into account the global difference of the distributions and is a good complement to the MK-S criterion.

For the third criterion for testing the fit of distributions, the likelihood function (LKV, Likelihood Value Test) was used as a measure of the fit of a probabilistic model to empirical data. The log value of the likelihood function (LKV) is calculated for empirical data [27, 14]. The likelihood function L depends on the random sample $T_1, T_2, \ldots, T_n$ and on parameters θ_j for which it takes maximum values. The general form of the likelihood function is given by the formula [33, 30]:

$$L\left(\theta_1, \theta_2, \ldots, \theta_k \middle| T_1, T_2, \ldots, T_n\right) = \prod_{i=1}^{n} f\left(T_i; \theta_1, \theta_2, \ldots, \theta_k\right) \tag{8}$$

where:

n is the number of failed components;
k is the number of parameters;
$\theta_j, j = 1,2,\ldots,k$ is the j-th parameter of the distribution;
$T_i, i = 1,2,\ldots,n$ is the time to failure of the i-th component.

In the case under consideration, the function was expanded to include factors taking account of right-censored data. The log-likelihood function is the sum of logarithms of probability density for particular lifetimes of the analysed component [18, 30]:

$$\Lambda\left(\theta_1, \theta_2, \ldots, \theta_k\right) = \ln L\left(\theta_1, \theta_2, \ldots, \theta_k \middle| T_1, T_2, \ldots, T_n\right) = \sum_{i=1}^{n} \ln f\left(T_i; \theta_1, \theta_2, \ldots, \theta_k\right) \tag{9}$$

where:

L is the likelihood function;
n is the number of failed components;
is the j-th parameter of the distribution;
$T_i, i = 1,2,\ldots,n$ is the time to failure of the i-th component.

Values of estimators of the unknown parameters $\theta_1, \theta_2, \ldots, \theta_k$ are determined by maximising the log-likelihood function $\Lambda\left(\theta_1, \theta_2, \ldots, \theta_k\right)$. A necessary condition for the existence of an extremum of this function is that all of its partial derivatives take the value 0.

To determine the estimators of the unknown parameters, partial derivatives $\frac{\partial \Lambda\left(\theta_1, \theta_2, \ldots, \theta_k\right)}{\partial \theta_j}$ of the function Λ are determined with respect to the parameters θ_j, $j = 1,2,\ldots,k$. To estimate the parameters, each partial derivative should be equated to zero and k equations should be solved:

$$\begin{gathered} \frac{\partial \Lambda\left(\theta_1, \theta_2, \ldots, \theta_k\right)}{\partial \theta_1} = 0 \\ \ldots\ldots\ldots \\ \frac{\partial \Lambda\left(\theta_1, \theta_2, \ldots, \theta_k\right)}{\partial \theta_k} = 0 \end{gathered} \tag{10}$$

In the last step, on the basis of each of the three goodness-of-fit criteria for all 11 distributions, ranks are assigned from the best-fitting to the worst-fitting theoretical distribution. Thus, the theoretical distributions are ordered separately for each criterion by assigning them successive natural numbers. Finally, based on the three rankings obtained and the weights assigned to the individual criteria, the aggregate criterion DESV is determined. This measure, for the *i*-th theoretical CDF (F_i), is given by formula (11):

$$\mathrm{DESV}(F_i) = \mathrm{RAVGOF}(F_i) \cdot \mathrm{WAVGOF} + \mathrm{RAVPLOT}(F_i) \cdot \mathrm{WAVPLOT} + \mathrm{RLKV}(F_i) \cdot \mathrm{WLKV} \tag{11}$$

where:

- $\mathrm{RAVGOF}(F_i)$ denotes the rank of the distribution F_i by the AVGOF criterion;
- $\mathrm{RAVPLOT}(F_i)$ denotes the rank of the distribution F_i by the AVPLOT criterion;
- $\mathrm{RLKV}(F_i)$ denotes the rank of the distribution F_i by the LKV criterion;
- WAVGOF denotes the weight of the AVGOF criterion;
- WAVPLOT denotes the weight of the AVPLOT criterion;
- WLKV denotes the weight of the LKV criterion.

The aggregate criterion DESV is therefore a weighted average of the individual ranks of theoretical distributions. After calculating the DESV value for the particular theoretical distributions, their final ranking is determined. The distribution with the lowest DESV value is identified as the best-fitting according to the aggregate criterion, and is assigned the number 1 in the ranking. The aggregate criterion is used to make the final selection of the distribution that best fits the empirical data among the theoretical distributions considered.

3. Subject of study

The aggregate criterion for ranking distributions of times to failure of selected vehicle components was applied based on operational data from a fleet of 45 urban rail vehicles of the same type, namely five-section low-floor Tramino S105P trams with total weight 42.5 tonnes and length approximately 32 metres. These are articulated, single-compartment vehicles. The tram can carry a maximum of 229 passengers, including 48 seated. The operational data covered the initial five years of use of the fleet, including two years covered by the warranty and three subsequent years under a maintenance contract [31, 9]. All trams were used in similar operating conditions, i.e. the same track infrastructure, similar daily and annual times of travel, and the same schedule and scope of (preventive) maintenance.

Fig. 2. Tramino S105P tram

From the database of failures in trams of the fleet under investigation, the lock of the driver's cab door was selected for testing of the time-to-failure distribution. This component failed 54 times during the first five years of operation, and generated 0.52% of all corrective maintenance costs [5]. The lock is mounted on the door between the passenger space and the driver's cab. To open the driver's cab door from the outside, the lock has to be opened mechanically with a special key. It was the bolting part of the lock that failed, becoming blocked and thus preventing the driver from opening the door and entering the cab. Depending on where the failure occurred, it was necessary to call the emergency maintenance service or to open the door using force, damaging the strike plate structure. On each such occasion the damaged lock was replaced with a new one. The cause of failure of the lock was excessive wear of the internal mechanism responsible for bolt extension, caused by a poorly selected construction material, as a result of which the lock stuck and sometimes prevented removal of the inserted key. A photograph of the lock is shown in Fig. 3.

Fig. 3. Driver's cab door lock

4. Empirical data

The process of tram operation is a valuable source of information serving to assess the required reliability parameters and to forecast maintenance costs. Operational information should be taken to include all data on events occurring during the operation and maintenance of trams [13]. These data play a key role in the planning and day-to-day management of vehicle fleet operation and maintenance, as well as in improving vehicle technology and construction [4, 35]. Operational information plays a particularly important role for operating companies, as it enables the proper planning of costs of operation, inspections and repairs, as well as assessment of the use of the means of transport [24, 34].

Before proceeding to the estimation of the parameters of probabilistic models of times to failure of selected vehicle components, the operational data should be appropriately prepared. For the investigated fleet of trams, operational data regarding individual vehicle components is right-censored of type I, which means that for a fixed period of use of the tram fleet the lock

Table 2. Right-censored times to failure of the lock in 5 years of operation

Dist. travelled [km]	F/S	Dist. travelled [km]	F/S	Dist. travelled [km]	F/S	Dist. travelled [km]	F/S
174,124	F	256,382	F	114,128	S	67,733	F
196,837	S	144,819	S	135,078	F	300,557	S
317,275	S	223,684	F	136,600	F	103,378	F
292,525	F	46,217	F	97,832	S	177,506	F
112,431	S	43,897	S	377,101	S	23,153	S
196,218	F	155,522	F	93,585	F	242,544	F
1,910	F	201,423	S	238,103	S	89,047	S
93,529	S	119,376	F	285,538	F	125,785	F
334,484	S	198,190	S	43,117	S	58,407	F
366,935	F	368,449	S	221,226	F	117,646	S
28,826	S	340,330	F	117,701	S	202,396	F
191,367	F	58,964	S	28,934	F	127,143	S
21,117	F	193,641	F	135,673	F	287,695	F
135,831	S	155,920	S	155,828	S	53,863	S
348,956	F	206,246	F	92,594	F	174,580	F
38,020	S	144,352	S	197,981	S	139,571	S
188,493	F	371,800	S	148,840	F	210,775	F
70,534	F	22,482	F	27,858	F	102,038	S
102,343	S	139,974	F	107,491	F	131,537	F
340,236	F	39,840	F	52,280	S	126,738	F
52,022	S	127,333	F	250,370	S	83,497	S
115,592	F	21,021	S	282,989	F	176,928	F
79,071	F	376,601	S	77,834	S	81,021	F
72,135	F	354,513	S	86,028	F	103,807	S
105,552	S	203,105	F	226,082	S	-	-

Table 3. Estimated parameters of tested distributions

1P-Exponential	**2P-Exponential**	**Normal**	**Lognormal**
$\hat{\lambda}$ = 3.413E-06	$\hat{\lambda}$ = 3.937E-06	$\hat{\mu}$ = 218,279.5	$\hat{\mu}'$ = 12.184
	$\hat{\gamma}$ = 21.117	$\hat{\sigma}$ = 115,461.8	$\hat{\sigma}'$ = 0.819
2P-Weibull	**3P-Weibull**	**Gamma**	**G-Gamma**
$\hat{\beta}$ = 1.745	$\hat{\beta}$ = 1.885	$\hat{\mu}$ = 11.53	$\hat{\mu}$ = 12.415
$\hat{\eta}$ = 255,316.9	$\hat{\eta}$ = 266,209.6	$\hat{\kappa}$ = 2.307	$\hat{\sigma}$ = 0.605
	$\hat{\gamma}$ = −10,300.12		$\hat{\lambda}$ = 0.857
Logistic	**Loglogistic**	**Gumbel**	
$\hat{\mu}$ = 211,755.9	$\hat{\mu}$ = 12.198	$\hat{\mu}$ = 274,770.6	
$\hat{\sigma}$ = 68,491.7	$\hat{\sigma}$ = 0.447	$\hat{\sigma}$ = 104,341.5	

F – failure, S – survival

failed and was replaced only in some vehicles, while in some vehicles it was replaced multiple times. Because the research concerns vehicles that are operated intensively, times to failure of individual components are expressed in kilometres. The time at which each vehicle comes into operation is known, and tram mileages at which components fail are recorded [33, 2]. The mileage of trams is used to determine the mileage of components at failure. The method of determining the mileage of failed vehicle components is presented in the paper [3].

Suitably prepared data are summarised in Table 2. They contain the exact time to failure of the tested component (the driver's cab lock) from the fleet of 45 trams under observation, expressed in kilometres and marked as F (failure), and the survival time of other locks that did not fail, marked as S (survival), also expressed as a number of kilometres travelled until the observations ended. At the time of the end of observations, the locks in all 45 vehicles were functional, although many of them had been replaced due to failure. Because the main reason for the replacement of locks is failure in the opening mechanism, all failures of this type were classified as mechanical failure.

Based on the data in Table 2, parameters were estimated for 11 theoretical distributions. The results of estimation for all examined distributions are given in tabular form (Table 3).

5. Identification of the best-fitting probability distribution

To select the best-fitting theoretical distribution out of the 11 considered, the aggregate ranking criterion described in section 2 was used. In determining the ranking of distributions, first the parameters of the theoretical distributions were estimated, and then the distributions were ranked based on the three criteria described. The results of this ranking procedure are summarised in Table 4.

The first column contains the name of the probability distribution. The second contains the values of the Kolmogorov–Smirnov AVGOF statistic – the probability of rejection of the working hypothesis for the MK-S statistic. The third column (AVPLOT) gives the mean absolute deviation of the theoretical CDF from the empirical CDF. The fourth column (LKV) gives the measures of goodness of fit determined using the log-likelihood criterion [8, 36, 21].

Table 4. Results of individual statistics for the data in Table 1

Distribution	AVGOF	AVPLOT	LKV
1P-Exponential	80.740	7.599	-720.16
2P-Exponential	55.253	5.666	-712.58
Normal	30.946	4.178	-715.65
Lognormal	14.276	2.790	-711.82
2P-Weibull	1.034	1.709	-709.71
3P-Weibull	2.396	1.870	-710.23
Gamma	0.045	1.585	-709.78
G-Gamma	**0.289**	**1.589**	**-709.66**
Logistic	36.386	3.730	-717.34
Loglogistic	1.716	1.768	-710.49
Gumbel	84.250	6.355	-723.73

After calculating the goodness-of-fit statistics for the three criteria and ranking the probability distributions, the next step was to assign weights to the criteria. In this study, the default values of weights selected by the software manufacturer were used. These are determined on the basis of engineering practice, resulting from many analyses conducted in industrial applications. Using the weights assigned to each criterion, the weighted average was calculated for the ranks obtained using the individual criteria. Finally, using the described DESV aggregate criterion, the final ranking of the eleven theoretical distributions was obtained. Weibull++ software was used for the estimation of parameters of the theoretical distributions and for constructing their rankings. For the analysed data, the following weights were assigned to the criteria: 40 for AVGOF, 10 for AVPLOT and 50 for LKV. After calculating the DESV value, the final ranking of distributions was determined (Table 5). The distribution with the lowest DESV value was identified as the best-fitting according to the aggregate criterion, and was assigned number 1 in the ranking. As shown in Table 5, the lowest value of the DESV statistic was obtained for the generalised gamma distribution. It was calculated from formula (11) as follows:

$$\mathrm{DESV} = (2\times 40)+(2\times 10)+(1\times 50)=150 \tag{12}$$

Thus, for the data contained in Table 2 regarding lock failures during five years of operation of the tram fleet, using the developed aggregate criterion, the generalised gamma distribution was identified as the best-fitting. This is reflected in the last column of Table 5.

Table 5. Weighted average values and ranking of distributions

Distribution	AVGOF	AVPLOT	LKV	DESV	Ranking
1P-Exponential	10	11	10	1010	9
2P-Exponential	9	9	7	800	7
Normal	7	8	8	760	6
Lognormal	6	6	6	600	5
2P-Weibull	3	3	2	250	3
3P-Weibull	5	5	4	450	4
Gamma	1	1	3	200	2
G-Gamma	**2**	**2**	**1**	**150**	**1**
Logistic	8	7	9	840	8
Loglogistic	4	4	5	450	4
Gumbel	11	10	11	1090	10

The estimated parameters μ, σ, λ for the reparameterised form of this distribution took the following values: $\hat{\mu}$=12.415; $\hat{\sigma}$=0.6058; $\hat{\lambda}$=0.8572. The calculated rate of failure of the lock was 0.000000617/km, and the average time to failure was 229,623 km.

To illustrate how the selected distribution matches the data, in Fig. 4 the data are presented on a probability plot of the generalised gamma distribution. The following figures show the reliability function (Fig. 5), the probability density function (Fig. 6) and a histogram of numbers of failures (Fig. 7).

In Fig. 4 the blue line represents the modelled probability of failure according to the generalised gamma distribution, and the red lines mark a two-sided 95% confidence interval. The reliability function graph (Fig. 5) shows the change in the reliability value over time, expressed as distance travelled in kilometres, indicating the trend in the behaviour of the tested component in terms of failures. The graph of the failure probability density function provides a visualisation of the distribution of data over time (Fig. 6). The histogram (Fig. 7) shows that a relatively large proportion of the failures occurred between 50,000 and 200,000 km.

The graphical presentation of the estimated functional characteristics (reliability, probability density) and the histogram of numbers of failures can be used to determine more easily the failure mode. This information is important when forecasting failures and determining the future cost of corrective maintenance resulting from them.

The presented analysis of the time to failure of the driver's cab lock shows that the best-fitting distribution, according to the aggregate criterion, is the generalised gamma distribution. It should also be noted that with successive failures, the aggregate method may indicate a different distribution as the best-fitting, because new data, especially if the quantity is large relative to that previously analysed, may follow a different model. In this situation, analysis of the plot of the probability distribution is very useful for pre-evaluating the fit of a selected theoretical model to the appropriate case.

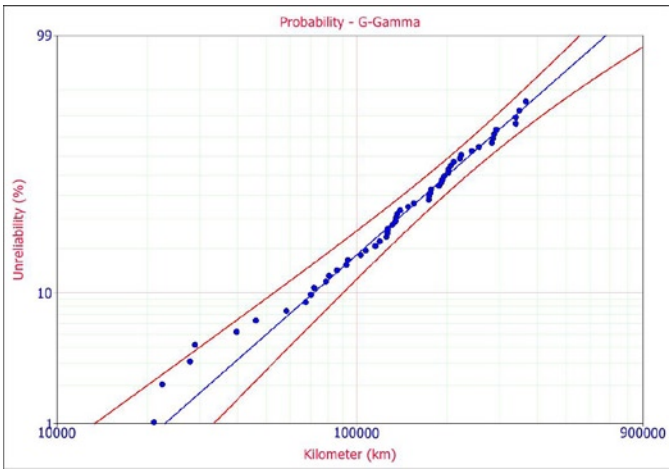


Fig. 4. Presentation of data on a probability plot of the generalised gamma distribution

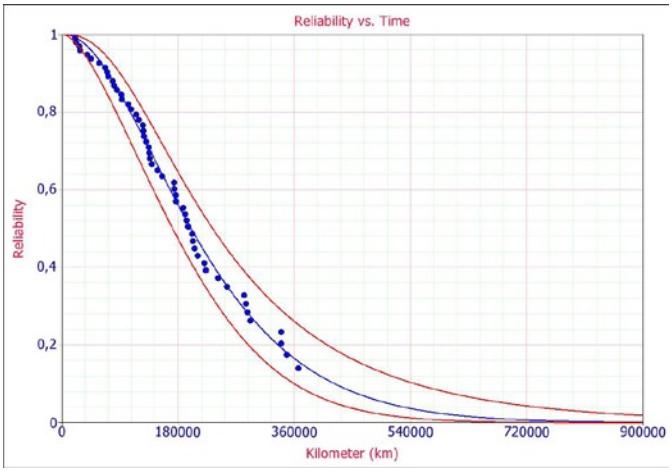


Fig. 5. Reliability function

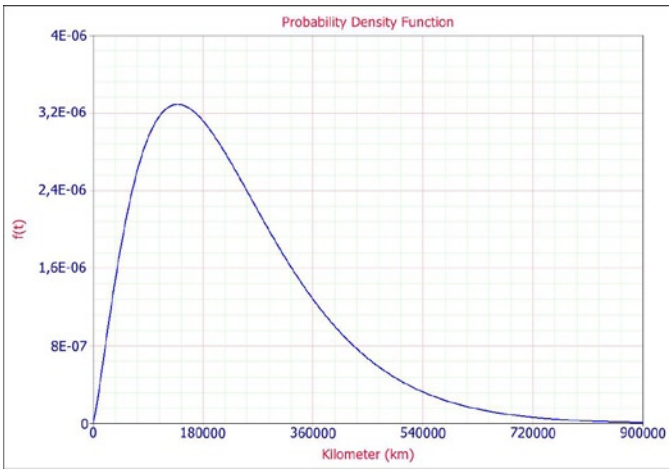


Fig. 6. Probability density function

When analysing failure data using an aggregate method, it has to be remembered that sometimes none of the statistical distribution models match the analysed data. In this case, the best of the worst solutions is obtained, which may poorly fit the data. In other cases, in which many models may be well matched to the empirical data, statistics alone are not enough; in such cases knowledge of the failure mechanism can be invaluable when selecting the most appropriate theoretical model. It is important to remember that, although the aggregate method applied to small samples will also rank selected

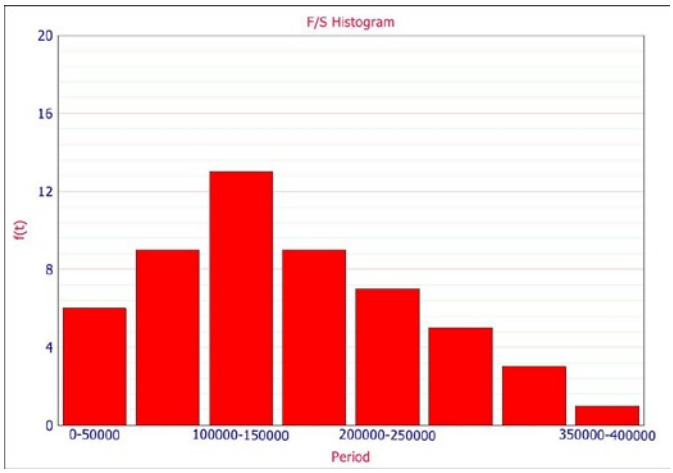


Fig. 7. Histogram

probability distributions depending on the number of parameters in a particular theoretical distribution, using it in such cases comes with a high level of uncertainty, and it is only recommended for use with larger data sets.

It should also be borne in mind that the two-parameter exponential distribution, the three-parameter Weibull distribution and the generalised gamma distribution contain a location parameter, a change in which causes a shift of the CDF and probability distribution function without changing their shape. On the other hand, the generalised gamma distribution is a complex model that can easily mimic many other distributions, and therefore often seems to be the best fitted to the analysed data.

When analysing data on probability plots, it can often be stated that they reflect more than one type of failure (e.g. fatigue, operational, construction, technological, etc.). In this case, all distributions ordered according to the aggregate selection method may turn out to be mismatched, because the developed method can only be used for a homogeneous type of failure of the examined component. In such situations, it is advisable to consider the possibility of using a mixture of distributions, e.g. a combination of two Weibull distributions.

6. Conclusions

The results obtained constitute an important argument for the possibility of using the proposed aggregate method of selecting a theoretical distribution for empirical data. By taking into account three criteria for assessing the accuracy of the fit, mistakes resulting from the use of only one of them can be avoided.

The use of only one criterion defining the quality of the fit of a theoretical to an empirical CDF may often prove insufficient, as it depends on many variables: mainly on the quantity of data and whether the data are full or censored, but primarily on the type of failure.

The aggregate method of identifying a theoretical distribution, taking into account three criteria, is a general method and has wide application, provided that the appropriate conditions are met: the number of observations must be large enough, and should contain accurate data on times to failure or to the end of observations. The modified K-S statistic (AVGOV) is sensitive to local deviations. On the other hand, the mean absolute deviation of the theoretical CDF from the empirical CDF (AVPLOT) is not so sensitive to local deviations; it takes into account the global difference of distributions, and is a good complement to the MK-S criterion. For the third criterion, the logarithm of the likelihood function (LKV), the size of the sample is important, because for small samples the value obtained may be strongly biased.

The benefits resulting from the correct selection of a random variable distribution for the time to failure of a renewable technical

object (a rail vehicle) are significant, among others due to the costs generated by failing to utilise fully the potential lifetime of the component, as well as losses resulting from unplanned corrective maintenance and vehicle downtime.

The research was financed from the following funds: PUT 04/43/SBAD/0111 and 05/51/SBAD/3584.

References

1. Abernethy R B. The New Weibull Handbook: Reliability & Statistical Analysis for Predicting Life, Safety, Survivability, Risk, Cost, and Warranty Claims (Fifth ed.), Florida, 2010.
2. Andrzejczak K, Selech J. Flexible Prediction of the Vehicle Component Damage. Transport Means 2018: Proceedings of the 22nd International Scientific Conference, Trakai, Lithuania, Part II, 2018; 987-990.
3. Andrzejczak K, Selech J. Generalised Gamma Distribution in the Corrective Maintenance Prediction of Homogeneous Vehicles. In: Kabashkin I, Yatskiv (Jackiva) I, Prentkovskis O. (eds) Reliability and Statistics in Transportation and Communication. RelStat 2018. Lecture Notes in Networks and Systems. Springer, Cham 2018; 68.4. Andrzejczak K, Selech J. Investigating the trends of average costs of corrective maintenance of public transport vehicles. Journal of KONBiN 2017; 41: 207-226, https://doi.org/10.1515/jok-2017-0011.
5. Andrzejczak K, Selech J. Quantile analysis of the operating costs of the public transport fleet. Transport Problems, 2017; 12 (3): 103-111.
6. Andrzejczak K. Statystyka elementarna z wykorzystaniem systemu Statgraphics [Elementary statistics using the Statgraphics system], Wyd. Politechniki Poznańskiej, Poznań 1997.
7. Bartnik G, Pieniak D, Niewczas A M, Marciniak A. Probabilistic model for flexural strength of dental composites used in modelling reliability of the "tooth-dental composite" system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (1): 136-141, https://doi.org/10.17531/ein.2016.1.18.
8. Bavuso S J. Aerospace Applications of Weibull and Monte Carlo Simulation with Importance Sampling, IEEE, Annual Reliability and Maintainability Symposium, Proc. 1997.
9. Dolce J E. Analytical Fleet Maintenance Management, SAE International, SUA, 1994.
10. Elmahdy E E. Modelling Reliability Data with Finite Weibull or Lognormal Mixture Distributions. Appl. Math. Inf. Sci. 2017; 11 (9), 1081-1089, https://doi.org/10.18576/amis/110414.
11. Ferreira L A, Silva J L. Parameter estimation for Weibull distribution with right censored data using EM algorithm. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (2): 310-315, https://doi.org/10.17531/ein.2017.2.20.
12. Fuc P, Rymaniak L, Ziolkowski A. The correlation of distribution of PM number emitted under actual conditions of operation by PC and HDV vehicles, WIT Transactions on Ecology and the Environment. WIT Press, 2013; 174: 207.13. Gill A. Optimisation of the technical object maintenance system taking account of risk analysis results. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (3): 420-431, https://doi.org/10.17531/ein.2017.3.13.
14. Hajkowski J, Popielarski P, Sika R. Prediction of HPDC casting properties made of AlSi9Cu3 alloy, Advances in Manufacturing, SPRINGER, Manufacturing 2017, 621-631, https://doi.org/10.1007/978-3-319-68619-6_59.
15. Hirose H. Bias Correction for the Maximum Likelihood Estimation in Two-parameter Weibull Distribution, IEEE Transactions on Dielectrics and Electrical Insulation 1999; 6: 1, https://doi.org/10.1109/94.752011.
16. https://www.reliasoft.com/Weibull [Accessed 2018].
17. Johnson R A, Miller I, Freund J E. Probability and Statistics for Engineers, eighth ed., Pearson Education Limited Co., UK, 2014.
18. Kececioglu D. Reliability & Life Testing Handbook, PrenticeHall, Inc., Englewood Cliffs, New Jersey, 1993; 1.
19. Lawless J F. Statistical Models and Methods for Lifetime Data, second ed., Wiley, 2002, https://doi.org/10.1002/9781118033005.
20. Lawless J F. Statistical Models And Methods for Lifetime Data, John Wiley & Sons, Inc., New York, 1982.
21. Lee E T, Wang J W. Statistical Methods for Survival. Data Analysis, John Wiley & Sons Inc; (3rd Edition), 2003, https://doi.org/10.1002/0471458546.
22. Legát V, Mošna F, Aleš Z, Jurča V. Preventive maintenance models - higher operational reliability. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (1): 134-141, https://doi.org/10.17531/ein.2017.1.19.
23. Liu J, Song B, and Zhang Y. Competing failure model for mechanical system with multiple functional failures. Advances in Mechanical Engineering 2018, 10(5) 1-16, https://doi.org/10.1177/1687814018773155.
24. Loska A. Exploitation assessment of selected technical objects using taxonomic methods, Eksploatacja i Niezawodnosc − Maintenance and Reliability 2013; 15, 1.
25. Młynarski S, Pilch R, Smolnik M, Szybka J. Methodology of network systems reliability assessment on the example of urban transport. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (2): 278-283, https://doi.org/10.17531/ein.2018.2.14.
26. Młyńczak M. Analiza danych eksploatacyjnych w badaniach niezawodności obiektów technicznych, Zeszyty Naukowe WSOWL, 2001; 1 (159).
27. Nelson W. Applied Life Data Analysis, John Wiley & Sons, Inc., New York, 1982, https://doi.org/10.1002/0471725234.
28. Perz P, Malujda I, Wilczyński D, Tarkowski P. Methods of controlling a hybrid positioning system using LabVIEW, 21th Scientific Polish-Slovak Conference "Machine Modeling and Simulations 2016", Procedia Engineering 2017; 177, 339-346, https://doi.org/10.1016/j.proeng.2017.02.235.
29. Pieniak D, Niewczas A M, Niewczas A, Bieniaś J. Analysis of Survival Probability and Reliability of the Tooth-composite Filling System. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2011; 2(50): 25-34.
30. ReliaSoft Corporation, Life Data (Weibull) Analysis Reference, ReliaSoft Publishing Tucson, AZ, 2008.
31. Research Project "Increase in the efficiency of functioning of public means of transport as a result of implementation of LCC and RAMS concepts in accordance with the IRIS standards based on integrated information technology system" financed by Polish National Center for Research and Development. No. PBS3/B6/30/2015.
32. Rojek I, Kujawińska A, Hamrol A, Rogalewicz M. Artificial neural networks as a means for making process control charts user friendly. In: Burduk A., Mazurkiewicz D. (eds.), Intelligent Systems in Production Engineering and Maintenance - ISPEM 2017, Advances in Intelligent

Systems and Computing, Springer, 637, 168-178, 2017, https://doi.org/10.1007/978-3-319-64465-3_17.
33. Selech J. Prognozowanie kosztów obsługiwania korekcyjnego pojazdów transportu masowego [Forecasting costs of corrective maintenance of mass transport vehicles]. Wydawnictwo Naukowe ITeE-PIB, Radom 2019, ISBN 978-83-7789-557-3.
34. Świderski A, Jóźwiak A, Jachimowski R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks, Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (2), 292-299, https://doi.org/10.17531/ein.2018.2.16.
35. Trojanowska J, Kolinski A, Galusik D, Varela M L R, Machado J. A methodology of improvement of manufacturing productivity through increasing operational efficiency of the production process. In: Hamrol A., Ciszak O., Legutko S., Jurczyk M. (eds) Advances in Manufacturing. Lecture Notes in Mechanical Engineering. Springer, Cham, 2018; 23-32, https://doi.org/10.1007/978-3-319-68619-6_3.
36. Waluś K J. Driver's Strategy and Braking Distance in Winter, Transport Means 2017: Proceedings of the 21st International Scientific Conference, Juodkrante, Lithuania. 2017; Part 2, 505 - 509, ISSN 1822-296 X, e-ISSN 2351-7034.
37. Wojtkowiak D, Talaśka K, Malujda I, Domek G. Estimation of the perforation force for polymer composite conveyor belts taking into consideration the shape of the piercing punch. The International Journal of Advanced Manufacturing Technology 2018, https://doi.org/10.1007/s00170-018-2381-3.
38. Ziółkowski J, Borucka A, Model Markowa w logistycznym zarządzaniu przedsiębiorstwem [Markov model in logistic management of enterprise], Journal of Konbin 2016; 2 (38), https://doi.org/10.1515/jok-2016-0027.
39. Żurek J, Ziółkowski J, Borucka A. Application of Markov processes to the method for analysis of combat vehicle operation in the aspect of their availability and readiness, Safety and Reliability - Theory and Applications - Čepin & Briš (Eds)©, Taylor & Francis Group, London, 2017; 2343-2352.

Jarosław SELECH
Karol ANDRZEJCZAK
Poznan University of Technology,
ul. Piotrowo 3A, 60-965 Poznań, Poland

E-mails: jaroslaw.selech@put.poznan.pl,
karol.andrzejczak@put.poznan.pl

Article citation info:
MI J, LI Y-F, BEER M, BROGGI M, CHENG Y. Importance measure of probabilistic common cause failures under system hybrid uncertainty based on bayesian network. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 112–120, http://dx.doi.org/10.17531/ein.2020.1.13.

Jinhua MI
Yan-Feng LI
Michael BEER
Matteo BROGGI
Yuhua CHENG

IMPORTANCE MEASURE OF PROBABILISTIC COMMON CAUSE FAILURES UNDER SYSTEM HYBRID UNCERTAINTY BASED ON BAYESIAN NETWORK

OPARTA NA SIECI BAYESOWSKIEJ MIARA WAŻNOŚCI PROBABILISTYCZNYCH USZKODZEŃ SPOWODOWANYCH WSPÓLNĄ PRZYCZYNĄ W WARUNKACH NIEPEWNOŚCI HYBRYDOWEJ SYSTEMU

When dealing with modern complex systems, the relationship existing between components can lead to the appearance of various dependencies between component failures, where multiple items of the system fail simultaneously in unpredictable fashions. These probabilistic common cause failures affect greatly the performance of these critical systems. In this paper a novel methodology is developed to quantify the importance of common cause failures when hybrid uncertainties are presented in systems. First, the probabilistic common cause failures are modeled with Bayesian networks and are incorporated into the system exploiting the α factor model. Then, probability-boxes (bound analysis method) are introduced to model the hybrid uncertainties and quantify the effect of uncertainties on system reliability. Furthermore, an extended Birnbaum importance measure is defined to identify the critical common cause failure events and coupling impact factors when uncertainties are expressed by probability-boxes. Finally, the effectiveness of the method is demonstrated through a numerical example.

***Keywords**: probabilistic common cause failure, Bayesian network, α factor model, extended Birnbaum importance.*

W przypadku nowoczesnych systemów złożonych, relacje zachodzące między komponentami mogą prowadzić do pojawienia się różnych zależności między ich uszkodzeniami, a tym samym do sytuacji w których kilka składowych systemu ulega uszkodzeniu jednocześnie w nieprzewidywalny sposób. Tego typu probabilistyczne uszkodzenia wywołane wspólną przyczyną (PCCF) mają ogromny wpływ na wydajność tych kluczowych systemów. W przedstawionym artykule opracowano nową metodę szacowania ważności PCFF w sytuacjach, gdy w systemie występują niepewności hybrydowe. W pierwszej kolejności, PCFF zamodelowano za pomocą sieci bayesowskich i włączono do systemu wykorzystującego model współczynnika α. Następnie, wprowadzono przedziały prawdopodobieństwa, tzw. probability boxes (bound analysis method), w celu zamodelowania niepewności hybrydowych i kwantyfikacji wpływu tych niepewności na niezawodność systemu. Ponadto zdefiniowano rozszerzoną miarę ważności Birnbauma, która pozwala zidentyfikować krytyczne zdarzenia PCCF oraz czynniki, które je wywołały, w przypadkach, gdy niepewności wyrażone są za pomocą probability boxes. Skuteczność metody wykazano na przykładzie numerycznym.

***Słowa kluczowe**: probabilistyczne uszkodzenie spowodowane wspólną przyczyną, sieć bayesowska, model współczynnika α, rozszerzona miara ważności Birnbauma.*

1. Introduction

The assessment of the reliability of modern industrial systems has to take into account various system characters since systems are becoming increasingly large and complex. As an example, reliability models have to take into account characteristics such as dynamic behavior [32], multiple failure mechanisms [10], components dependent relationships, uncertainties, etc. [15, 18]. Several conventional combinatorial methods have been developed and proved to be effective tools for system reliability modelling and assessment, including reliability block diagram (RBD), fault tree analysis (FTA) [12], Markov chains, Bayesian network [11, 23], etc. Nevertheless, when considering model performance, computational efficiency and executive complexity, these traditional models present advantages, but also serious limitations [3]: (i) Static fault tree and RBD model can map system components to events, but fail to capture the dynamic behavior; dynamic fault tree models are needed to model time-dependent behaviors, increasing significantly the complexity of the investigated model. (ii) Markov chain can deal with dynamic behavior, but it is limited to exponential distribution for failure behaviors. Markov chain is also faced with state space exponential explosion problems when applied to systems of large size, Because Markov chain method consider all relationships among parent nodes, children nodes, and even sharing nodes. (iii) Due to conditional independence assumptions between the random variables and dependence separation among the nodes in Bayesian network, a child node in a Bayesian network is only affected by a limited number of parent nodes [11, 12]. (iv) Otherwise, Bayesian networks provide a powerful capability of probability reasoning, dynamic behavior modeling and multi-model synthesis [23]. These advantages prompt Bayesian network to be a widely used method in

reliability modelling and assessment of a diversity of large engineering systems.

In engineering practice, unavoidable uncertainties are of uttermost importance for system reliability assessment. The combination of both aleatory (stochastic) uncertainty and epistemic (lack of knowledge) uncertainty leads to the framework called "mixed uncertainty" or "hybrid uncertainty", and it is ubiquitous in engineering systems [4]. Uncertainties mainly arise from the following aspects: observational uncertainty, model uncertainty and parametric uncertainty. The purpose of uncertainty analysis is developing advanced approaches to reduce those uncertainties, thus leading to more accurate analysis and assessment of system reliability [29]. The uncertainty characterization models can be divided into three types: classic probabilistic analysis, non-probabilistic models, and imprecise probability model. Imprecise probability models, including evidence theory [7, 14], probability-box (p-box) theory [8], fuzzy probability theory [25, 28], etc. have been proved to be more appropriate for hybrid uncertainty. In particular, the essence of p-box theory is the combination of classic probability theory and interval arithmetic, it's a very effective tool to treat imprecise probabilities, allowing for the comprehensive propagation of hybrid uncertainty [6, 24].

As defined by the Nuclear Energy Agency (NEA), common cause failures (CCFs) are the simultaneous failure events of two or more components in the same common cause component group (CCCG). CCFs are caused by shared initiating events which also called "coupling impact factors" [30]. CCFs directly connect the failure event with root causes; thus, research work on CCFs can build the cause-effect relationship between components failures and failure causes. Multiple of parametric models, such as β-factor model and α-factor models, generically classified as "ratio models" have been developed for quantification of CCFs. In addition, additional CCF models allow for direct representation of the CCF events, such as the square-root method; and shock models, such as binomial failure rate model, have also been proposed. For reliability analysis and assessment of system with CCFs, these parametric models have been extended to incorporate CCFs into system fault tree model [13, 16], Bayesian network model [17], etc. These methods are especially widely used in probabilistic safety assessment of large complex systems with high reliability and long lifetime requirements [9].

Probabilistic CCF (PCCF) is a generalized model of CCF that can characterize the simultaneous failures of components in CCCGs with different probability of occurrence. When employing ratio models for PCCF analysis, even if the β-factor model is the most widely used method thanks to its simplicity, α-factor model is receiving increased interest as well since it can model multiplicities of CCFs and can build a bridge between failure events and coupling causes [31]. When dealing with the reliability assessment of systems taking into consideration PCCFs by means of static fault tree model, some explicit and implicit modelling methods were proposed by Wang, et al. [26] to estimate the reliability of systems with arbitrary components types and different component failure distributions. Thereafter, Wang, et al. [27] extended these models and proposed both an explicit and implicit method to analyze reliability of phased-mission systems with PCCFs. Zhu, et al. [33] proposed a stochastic computational approach to deal with the reliability overestimate of dynamic fault trees with PCCFs when dynamic behaviors are considered in redundant system. Additionally, when epistemic uncertainty are also present in systems, Zuo et al. [34] evaluated the system reliability when PCCFs are specified as interval value based on evidential network, and the Birnbaum importance was extended to measure the contribution of components to system reliability. Based on evidential network, Qiu et al. [21] proposed a valuation-based system method for system reliability analysis with consideration of parametric uncertainty and CCFs. Mi, et al. [17] incorporated CCFs and uncertainties into evidential network to study the reliability of multi-state systems. These methods mainly focus on epistemic uncertainty and CCFs, and it needs to be emphasized that there still lack of research work on system reliability when hybrid uncertainties and PCCFs are both considered.

Hence, in recent years, the research works of CCF are mainly focusing on the quantification models and method extensions, while few works are carried on estimation of the importance measure of different types of CCFs and of impact factors to system reliability [1, 22], especially when aleatory uncertainty and epistemic uncertainty are unavoidably, present in systems. This research gap will be of considerable significance to safety critical industrial areas seriously affected by CCFs, such as aerospace and nuclear industry. The ranking of CCF events and impact factors can give explicit guidance for system renew design, and also meaningful for maintenance measure formulating and fault eliminating.

To evaluate the importance of various CCF events to system reliability, the CCF events should be modelled and expressed in system reliability model. Therefore, this paper is organized as follows, firstly, the CCF events are modelled by Bayesian network based on alpha factor model, and incorporated into system Bayesian network. Then the hybrid uncertainties are expressed by p-boxes, and the Birnbaum importance are extended as EBI (extended Birnbaum importance) which can be used to define the importance of CCF events to system reliability. Finally, a numerical example is used to realize the impact measure of CCF to system reliability.

2. Common Cause Failure Modeling Based on Bayesian Network

2.1. Parametric model for common cause failure

α-factor model is a multiparameter method which can be used to quantify all kinds of CCFs. The definition of the α-factor, indicated as α_k, is the fraction of the total failure probability of events that occur in the system and involve the failure of k components caused by a common cause. For a common cause component group (CCCG) with m components in the same type, which also called a CCCG of size m, the sum of all α_k equals to 1. After a series of experiments, the number of basic failure events is collected, and the number of failure events with k components failure based on a common cause is n_k which can be computed by weighted impact vector method [19, 20]. Then the alpha factor can be estimated by using the maximum likelihood estimator when there are sufficient test data, and:

$$\hat{\alpha}_k = \frac{n_k}{\sum_{k=1}^{m} n_k} \tag{1}$$

Then, a common cause vector $\boldsymbol{\alpha}_{CCCG}^{m}$ is defined to represent the effect degree of different CCF events on failure of each component in a CCCG of size m and:

$$\boldsymbol{\alpha}_{CCCG}^{m} = \left[\alpha_1, ..., \alpha_i, ..., \alpha_m\right] \tag{2}$$

Besides, since β factor model can only get an approximate scope by engineering experiences, but α factor model has the ability of integrating experts' judgments of system and past data, which makes it to be a more suitable parameter model in practice engineering. When the total failure probability of a component is P_t, the occurrence probability of the corresponding failure events for staggered testing is given by:

$$P_k^m = \frac{\alpha_k}{\binom{m-1}{k-1}} P_t \tag{3}$$

When the occurrence probability of component P_t is specified, alpha factors can be calculated by Eqs. (1) and (2), then the occurrence probabilities of different CCF scenarios can be calculated by Eq. (3).

2.2. Bayesian network modeling of component in common cause component groups

Bayesian networks are a widely used methodology for reliability modelling; they are composed of nodes, which represent binary state random variables, and edges, which represent dependencies between these variables, respectively. In a Bayesian network with n nodes, where X_i is the corresponding random variable of node i, for any X_i (i=1,…,n), there exists $\pi(X_i) \subseteq \{X_1,...,X_{i-1}\}$ which causes variable X_i to be conditional independent from all the variables in the set $\{X_1,...,X_{i-1}\}$. Thus, based on the chain rule and conditional independence of Bayesian network, the joint distribution of n variables can be derived by the following formula as:

$$P(X_1,...,X_n) = \prod_{i=1}^{n} P(X_i | X_1,...,X_{i-1}) = \prod_{i=1}^{n} P(X_i | \pi(X_i)) \quad (4)$$

When $\pi(X_i) = \varnothing$, the conditional distribution $P(X_i|\pi(X_i))$ will degrade into the marginal distribution $P(X_i)$. Thanks to the decomposition of joint distribution, it is possible to greatly reduce the complexity of a Bayesian network model.

In this framework, the failure event of a component in a CCCG can be decomposed based on Eq. (3). For example, for a component X in a CCCG with 3 components, the failure event of a component, e.g. the failure event identified by the random variable X_1, can be further characterized as independent failure, $X_{1\text{-ind}}$, two components CCF, X_{12} and X_{13}, and three components CCF, X_{123}. Then, the Bayesian network of the failure of the component can be built as shown in Fig. 1. Based on the conditional independence between variables and reasoning mechanism of Bayesian network, the probability of component X_1 is:

$$\begin{aligned} P(X_1) &= \sum_{X_{1\text{-}ind}, X_{12}, X_{13}, X_{123}} P(X_{1\text{-}ind}, X_{12}, X_{13}, X_{123}, X) \\ &= \sum_{X_{1\text{-}ind}, X_{12}, X_{13}, X_{123}} \prod P(X_i, \pi(X_j) = \varnothing) \end{aligned} \quad (5)$$

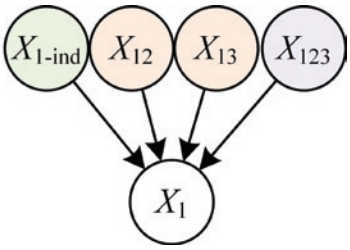


Fig. 1. BN of decomposed component in CCCG of size 3

When an arbitrary independent failure event or CCF event occurs, the failure of component will be triggered. By using 0 and 1 to represent the failure and functioning states of X, respectively, the conditional probability table (CPT) of node X is listed in Table 1.

In the case when 3 components of the same type are connected in a parallel system, the failure of each component within the CCF events is represented through a Bayesian network shown in Fig. 1, and the Bayesian network of the whole system can be further assembled and it is shown in Fig. 2. Finally, the probability of the system can be evaluated by exploiting the forward reasoning of Bayesian network and Eq. (4), and expressed as:

$$\begin{aligned} P(S) &= \sum P(S, X_1, X_2, X_3, X_{1\text{-ind}}, X_{2\text{-ind}}, X_{3\text{-ind}}, X_{12}, X_{13}, X_{23}, X_{123}) \\ &= \sum P(S|X_1, X_2, X_3) \cdot P(X_1|X_{1\text{-ind}}, X_{12}, X_{13}, X_{123}) \\ &\cdot P(X_2|X_{3\text{-ind}}, X_{12}, X_{23}, X_{123}) \cdot P(X_3|X_{3\text{-ind}}, X_{13}, X_{23}, X_{123}) \\ &\cdot P(X_{1\text{-ind}}) \cdot P(X_{2\text{-ind}}) \cdot P(X_{3\text{-ind}}) \cdot P(X_{12}) \cdot P(X_{13}) \cdot P(X_{23}) \cdot P(X_{123}) \\ &= \sum \left(P(S|\pi(S)) \cdot \prod_{i=1}^{3} P(X_i|\pi(X_i)) \prod P(X_j, \pi(X_j) = \varnothing) \right) \end{aligned} \quad (6)$$

When the CPTs and marginal distributions are given, the system reliability can be further evaluated.

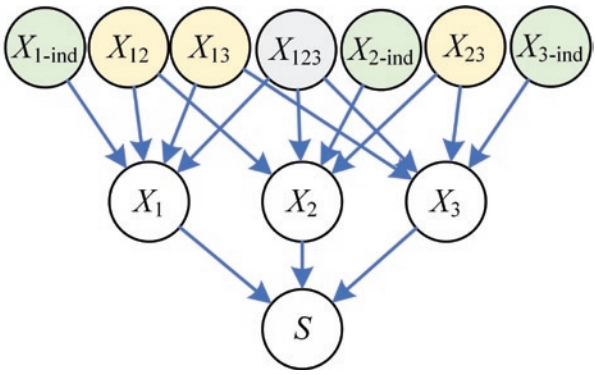


Fig. 2. BN of 3 component parallel system with CCFs

3. Bayesian Network Reasoning and Extended Importance Measure under Hybrid Uncertainties

3.1. Probability-box for uncertainty expression

The term "hybrid uncertainties" is used to identify uncertainty quantification and propagation procedures that include both aleatory uncertainty and epistemic uncertainty. The aleatory uncertainty characterizes the inherent randomness typical of some physical processes which cannot be eliminated or reduced and is quantified by means of probability theory. On the other hand, epistemic uncertainty is present in systems due to lack of knowledge, insufficient data, etc., but it can be reduced by providing more data and increasing the knowledge of the system. P-box theory has been proved to be an effective method

Table 1. CPT of decomposed component

X_{ind}	X_{12}	X_{23}	X_{123}	X	
				0	1
0	0	0	0	1	0
0	0	0	1	1	0
...	...	...	...	...	...
1	1	1	1	0	1

to analyze aleatory and epistemic uncertainty in systems. The probability expression of p-box boundaries includes the aleatory uncertain information of system performance, while the area between the upper and lower bounds represent the epistemic uncertain information.

For a random variable X affected by hybrid uncertainty, its probability distribution is not identified by a unique cumulative distribution $F_X(t)$, but by an upper and lower bound, consisting of a p-box $[\underline{F}_X(t), \overline{F}_X(t)]$. The overall slant of a p-box represents aleatory uncertainties; while the epistemic uncertainty is represented by the breadth between the upper and lower bound of the p-box. Based on this definition, the p-box is extended and used in system reliability analysis, sensitivity analysis, risk analysis, etc.

As an example of p-box appearing in the estimation of the reliability of a system, let's consider the case when the lifetime of a component is assumed to follow a Weibull distribution. The shape parameter β and scale parameter η are affected by imprecise information and defined as interval parameters, which varies in $[\underline{\beta}, \overline{\beta}]$ and $[\underline{\eta}, \overline{\eta}]$, respectively.The lifetime of a component is described by a non-negative random variable X_i on the real number $\mathbb{R}$, and $\overline{R}_i(t)$ and $\underline{R}_i(t)$ are the bounding reliability functions for random variable X, and $R_i(t) = P\{X_i > t\}$. Then the p-box variable which is employed to express the system reliability can be defined as [17]:

$$\Re_i = \left\{ R_i(t), \forall t \in \mathbb{R} \middle| \underline{R}_i(t) \le R_i(t) \le \overline{R}_i(t) \right\} \tag{7}$$

For example, for a component with lifetime distribution follows Weibull distribution where the scale parameter is within the interval [1.68, 1.86] and the shape parameter is within interval [2.08, 2.32], the Weibull p-box is constructed by taking the envelopes of those distributions and is shown in Fig. 3.

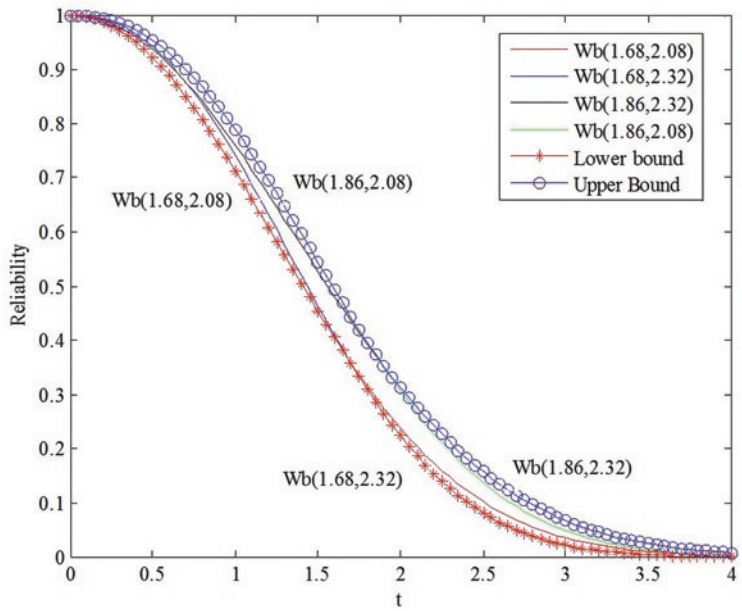


Fig. 3. Example reliability p-box with Weibull distribution

3.2. Bayesian network reasoning with hybrid uncertainty

For a Bayesian network with n+m+1 nodes, the variables of n root nodes are represented as x_i (i=1,2,…,n), the variables correspond to intermediate nodes is y_j (j=1,2,…,m), the leaf node variable is T, based on the chain rule of Bayesian network, the system reliability can be calculated by the following equation

$$\begin{aligned} R_S &= P(T=0) \\ &= \sum_{x_1,\ldots,x_n,y_1,\ldots,y_m} P(x_1,\ldots,x_n,y_1,\ldots,y_m,T=0) \\ &= \sum_{\pi(T)} P(T=0|\pi(T)) \prod_{j=1}^{m} P(y_j|\pi(y_j)) \prod_{i=1}^{n} P(x_i) \\ &= \sum_{\pi(T)} P(T=0|\pi(T)) \sum_{\pi(y_1)} P(y_1|\pi(y_1)) \times \ldots \\ &\times \sum_{\pi(y_m)} P(y_m|\pi(y_m)) \times \ldots \times R(x_1) \times \ldots \times R(x_n) \end{aligned} \tag{8}$$

where $\pi(T)$ is the parent of leaf node T, and $\pi(y_m)$ represents the parent of intermediate node y_m. When hybrid uncertainties are taken into account in the system, and the reliability of basic components is expressed by p-boxes in Eq. (7), then, the occurrence probability of decomposed independent failure events and CCF events can be obtained by Eq. (3) when the alpha factors are given or estimated. Since the reliability function is a monotone decreasing function, the p-box of the system reliability can be further derived and expressed as:

$$\begin{aligned} [R]_S &= [P](T=0) \\ &= \sum_{x_1,\ldots,x_n,y_1,\ldots,y_m} [P](x_1,\ldots,x_n,y_1,\ldots,y_m,T=0) \\ &= \sum_{\pi(T)} P(T=0|\pi(T)) \prod_{j=1}^{m} P(y_j|\pi(y_j)) \prod_{i=1}^{n} [P](x_i) \\ &= \sum_{\pi(T)} P(T=0|\pi(T)) \sum_{\pi(y_1)} P(y_1|\pi(y_1)) \times \ldots \\ &\times \sum_{\pi(y_m)} P(y_m|\pi(y_m)) \times \ldots \times [R](x_1) \times \ldots \times [R](x_n) \end{aligned} \tag{9}$$

After the upper and lower bound of system reliability p-box are obtained, the hybrid uncertainty of system can be intuitively shown.

3.3. Extended Birnbaum importance measures

The Birnbaum importance measure has been commonly used to evaluate the contributions of components to the reliability of binary coherent system. In this paper we focus on the dependent failure especially CCFs. The failure events of the components are originally not independent; however, after the decomposition of the component failure events by partial alpha-factor model, the decomposed basic events are independent [2]. The Birnbaum importance measure can be further extended and employed to evaluate the contributions of CCF events and independent events to the system reliability.

For a Bayesian network which has n root nodes and composed of both CCF events and independent failure events, and where $n = \sum_{k=1}^{K} m_k$, the system reliability is represented as $R_S(p_{x_1},\ldots,p_{x_j},\ldots,p_{x_n})$. In addition, the state of component or event j, indicated by x_j, is considered. x_j = 0 represents the j-th component in a false state, while x_j = 1 represents the event j as true. By explicitly assigning these two possible states to the j-th component, the reliability of the system at where x_j is either 0 or 1 is expressed as $R_S(p_{x_1},\ldots,p_{x_{j-1}},0,p_{x_{j+1}},\ldots,p_{x_n})$ and $R_S(p_{x_1},\ldots,p_{x_{j-1}},1,p_{x_{j+1}},\ldots,p_{x_n})$, respectively. Then, by extending the basic definition of BI, which is defined as the partial derivative of the system reliability function, the extended Birnbaum importance (EBI) of the event j is defined as:

$$EBI_j = \frac{\partial R_S\left(p_{x_1},...,p_{x_n}\right)}{\partial p_{x_j}}$$
$$= R_S\left(p_{x_1},...,p_{x_{j-1}},1,p_{x_{j+1}},...,p_{x_n}\right) - R_S\left(p_{x_1},...,p_{x_{j-1}},0,p_{x_{j+1}},...,p_{x_n}\right) \quad (10)$$

where p_{x_j} is the occurrence probability of event *j*. The EBI measure can be used to rank the importance of failure event i when aleatory uncertainties are exclusively considered. However, when epistemic uncertainty is also present in the system, and the reliability of the components is represented by p-boxes (refer to Section 3.1), then p-box which used to express the extended BI of event *j* can be further obtained and calculated by:

$$EBI_j^I = \frac{\partial \Re_S}{\partial p_{x_j}} = \left\{EBI_j(t), \forall t \in \mathbb{R} \middle| EBI_j^L(t) \le EBI_{ji}(t) \le EBI_j^U(t)\right\} \quad (11)$$

where the lower bound and upper bound of EBI can be further computed by the following two global optimization algorithms:

$$EBI_j^L: \min\ R_S\left(p_{x_1},...,p_{x_{j-1}},1,p_{x_{j+1}},...,p_{x_n}\right) - R_S\left(p_{x_1},...,p_{x_{j-1}},0,p_{x_{j+1}},...,p_{x_n}\right)$$
$$EBI_j^U: \max\ R_S\left(p_{x_1},...,p_{x_{j-1}},1,p_{x_{j+1}},...,p_{x_n}\right) - R_S\left(p_{x_1},...,p_{x_{j-1}},0,p_{x_{j+1}},...,p_{x_n}\right)$$
$$\text{s.t.}\quad \underline{R}_S\left(p_{x_j}=1\right) \le R_S\left(p_{x_1},...,p_{x_{j-1}},1,p_{x_{j+1}},...,p_{x_n}\right) \le \overline{R}_S\left(p_{x_j}=1\right)$$
$$\underline{R}_S\left(p_{x_j}=0\right) \le R_S\left(p_{x_1},...,p_{x_{j-1}},0,p_{x_{j+1}},...,p_{x_n}\right) \le \overline{R}_S\left(p_{x_j}=0\right) \quad (12)$$

After the BIs of varieties of CCF events are obtained and expressed by p-boxes, the interval-valued BI at specific time *t* can also be derived, then the ranking of the contribution of CCF events to system reliability can be further obtained.

4. Case Study

4.1. System description

An arbitrary 13-component non-repairable complex system in Ref. [5] is used to illustrate the proposed method in this section. The compound system is further transformed into series-parallel system as shown in Fig. 4. All 13 components can be divided into 5 groups, and the components in the same group have the same lifetime distribution. The components classification and corresponding hypothetical lifetime distributions are defined in Table 2.

4.2. Reliability analysis of the system

The reliability modeling and analysis of the example system is carried out with the following steps:

Step 1: Develop the Bayesian network model of the system without considering CCFs and hybrid uncertainty. Based on the system

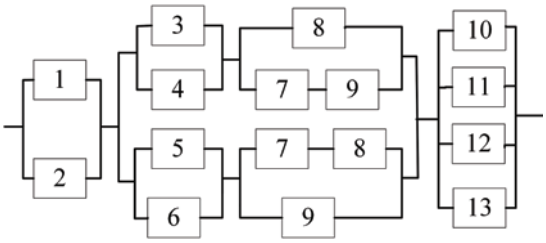


Fig. 4. The structure of example complex system

structure in Fig. 4, the Bayesian network of the system is constructed, as shown in Fig. 5. The nodes 1 to 13 correspond to the basic failure events of 13 basic components, while the nodes 14-15 are the intermediate nodes and finally node 25 is the leaf node, used to represent the event of system failure. The CPTs of a simple 3 nodes Bayesian network structure under series and parallel system structure are shown in Table 3, and the CPTs of a Bayesian network structure with more nodes can be easily inferred based on the system structure and failure mechanisms., The system reliability can be obtained by means of inference of the Bayesian network, and it is shown in Fig. 6. In order to validate the proposed approach, the system reliability from Bayesian network is compared the reliability computed by survival signature; the two results agree very well, proves the validity of the proposed method.

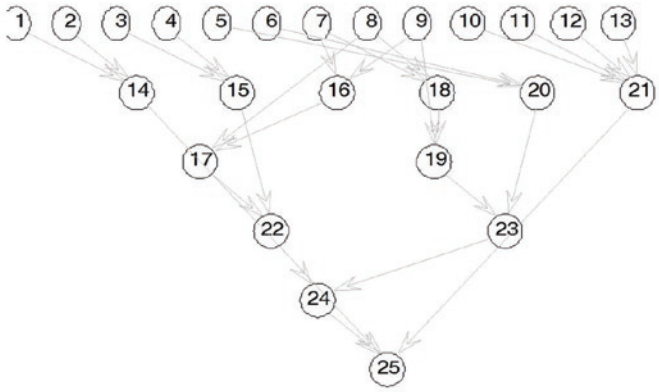


Fig. 5 System Bayesian network without CCFs

Step 2: System reliability without CCF events but with hybrid uncertainties. When aleatory uncertainty and epistemic uncertainty are considered in this system, the system Bayesian network is represented by the same structure as shown in Fig. 5. The upper bound and lower bound of system reliability can be calculated based on Eqs. (10) to (12), and are shown in Fig. 7. By contrast, the results obtained by survival signature method within those two conditions: (1) with consideration of hybrid uncertainties; (2) without hybrid uncertainties. The results are also shown in Fig. 7. It shows that regardless of whether or not the hybrid uncertainties are considered, the system reliability computed by survival signature and traditional Bayesian network is with consistent, and these results are within the upper and lower bounds which calculated by p-box method. The

Table 2. Parameters of system components

Comp. type	Dist. type	Dist. parameters	Imprecise Dist. parameters	CCF parameters
1 (1, 2)	Wb	(1.8,2.2)	([1.68,1.86], [2.08,2.32])	{0.95,0.05}
2 (3-6)	Exp	1.2	[1.07,1.33]	{0.8,0.1,0.05,0.05}
3 (7)	Wb	(2.3,1.6)	([2.12,2.51], [1.38,1.72])	{1}
4 (8, 9)	Wb	(3.2,2.6)	([2.99,3.41], [2.51,2.97])	{0.9,0.1}
5(10-13)	Exp	2.1	[2.01,2.28]	{0.75,0.1,0.1,0.05}

Table 3. The CPTs of 3 nodes Bayesian network under series and parallel system structure

X_1	X_2	Y (Series)		X_1	X_2	Y (Parallel)	
		0	1			0	1
0	0	1	0	0	0	1	0
0	1	1	0	0	1	0	1
1	0	1	0	1	0	0	1
1	1	0	1	1	1	0	1

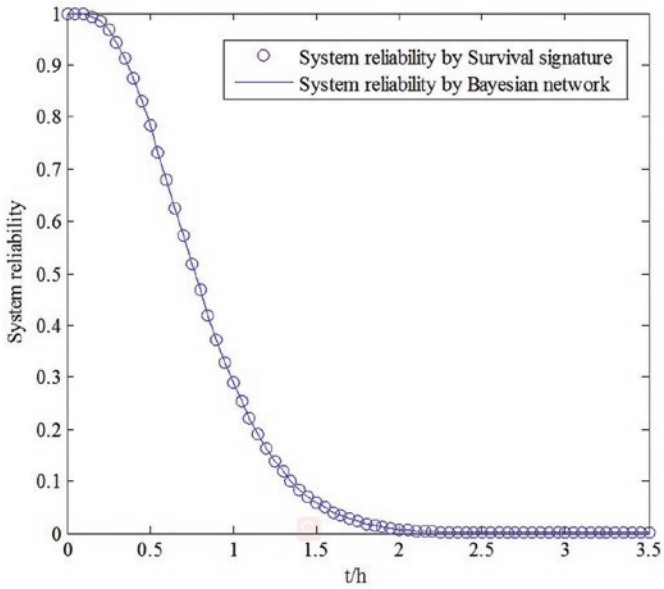


Fig. 6 System reliability ignore uncertainties and CCFs

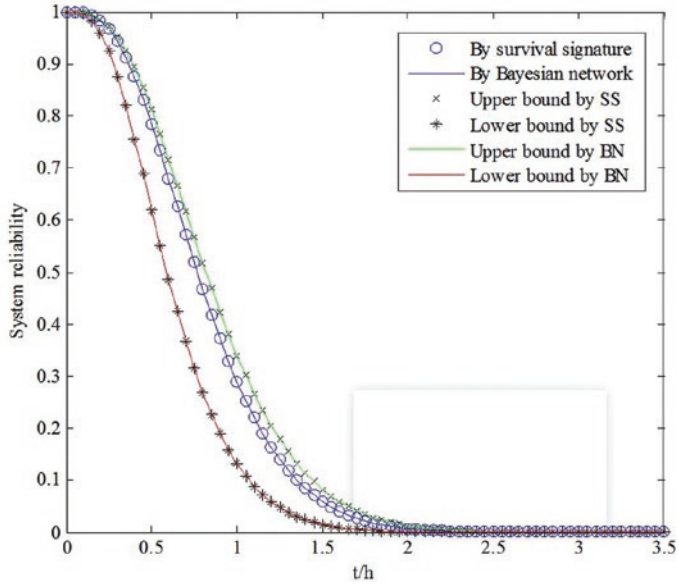


Fig. 7. System reliability with hybrid uncertainties

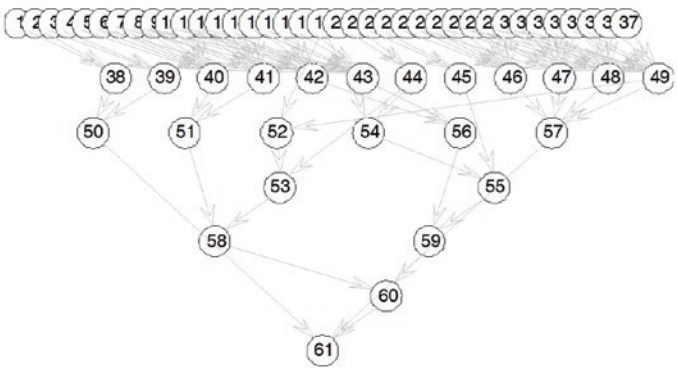


Fig. 8. System Bayesian network with CCFs

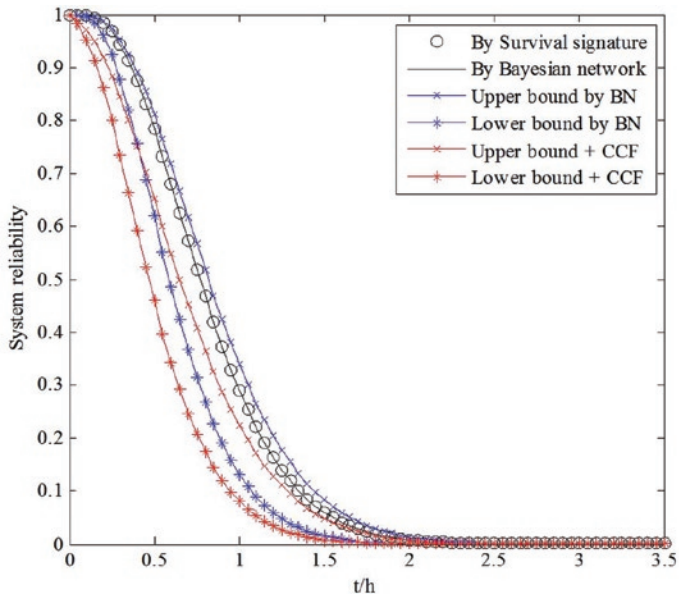


Fig. 9. System reliability with CCF and hybrid uncertainties

trend of the system reliability contains the aleatory uncertain information of system, and the breadth of reliability p-box can clearly reveal the epistemic uncertainty of system.

Step 3: Develop the Bayesian network model of system with CCF events. Based on the CCF modeling method illustrated in Section 2, a new Bayesian network model of this example system can be built and shown in Fig. 8. Where the nodes 1 to 37 are independent failure events of components and various different CCF seniors of CCFGs, nodes 38 to 49 correspond the basic failure events of 13 components, nodes 50 to 60 are the intermedia nodes, and the system failure events is represented by node 61. The CPTs of nodes in components layer are similar to Table 1.

Based on the partial alpha factor model, the probabilities of independent failures and various CCF events can be calculated by Eqs. (1) to (3). Then system reliability can be further computed and shown in Fig. 9. The reliability of the system is decreased with a more serious tendency when CCFs are considered, which further shows the significant influence of CCF events to system reliability.

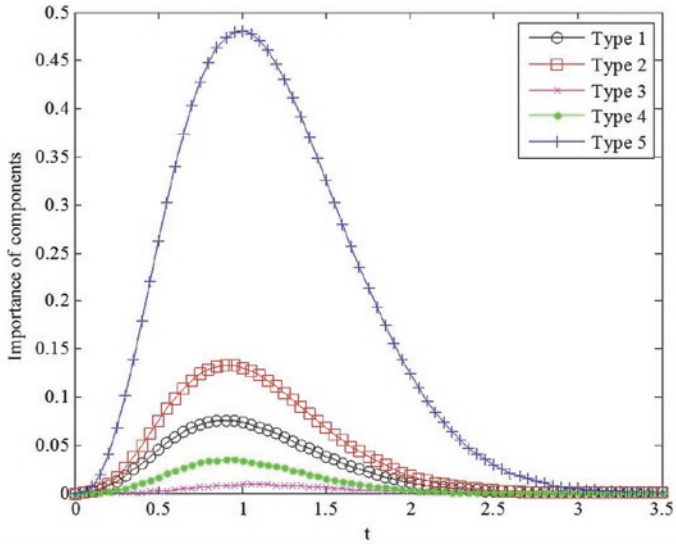


Fig. 10. Importance of different types of components

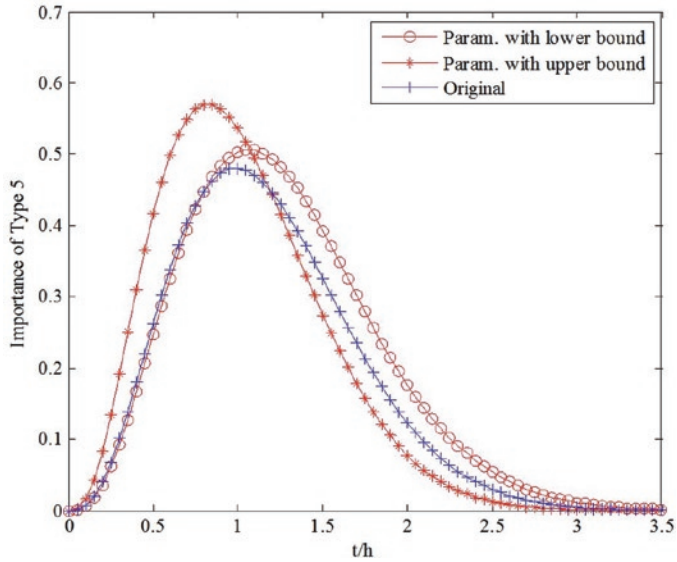


Fig. 11. The importance changes of component type 5 with uncertainties

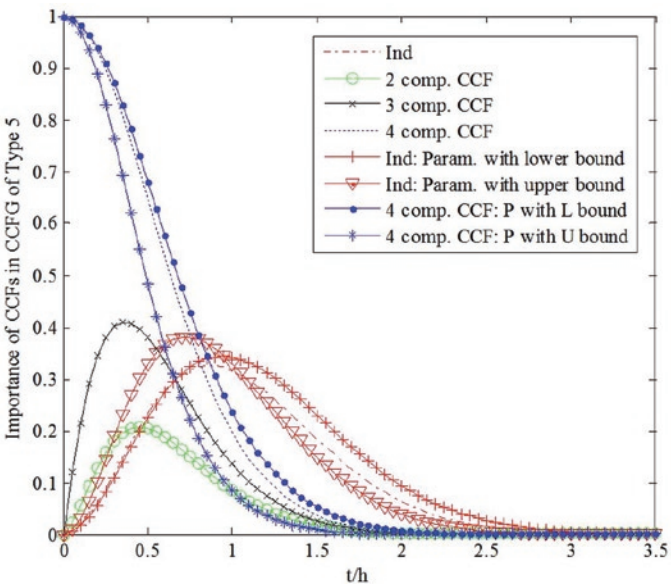


Fig. 12. The EBI of various CCFs of component type 5

4.3. Importance analysis of components and common cause failures

Based on Eqs. (10) to (12), the importance of all types of components without considering CCFs can be calculated and shown in Fig. 10. Then the ranking of component importance is *RI*(Type5)>*RI*(Type2)>*RI*(Type1)>*RI*(Type4)>*RI*(Type3). In order to indicate the effect of uncertainties to component importance, the importance of component type 5 with uncertainties is shown in Fig. 11. There is a big difference of type 5 after considering uncertainties, and there is an intersection between the lower and upper bound of importance.

To indicate the importance of CCFs to system, based on the definition in Section 3.3, as shown in Fig. 12, the EBIs of various CCFs of component type 5 are computed. Then the importance ranking of various CCF events can be gotten at any specific time. Furthermore, after getting the impact factors of CCFs, it will give a more accurate guide for system design and maintenance measure making.

5. Conclusions

This paper mainly discusses the importance measure of CCF events to system reliability with consideration of hybrid uncertainties. Firstly, considering the existing theory research of CCF and system modeling, we primarily model the system reliability by Bayesian network, and then the CCFs are incorporated into system Bayesian network model based on alpha factor model. The result compared with survival signature shows the validity of Bayesian network model and CCFs can decrease system reliability with a heavy tendency. When considering hybrid uncertainties in system, we extend the Birnbaum importance for CCF events on the basis of p-box and get the upper and lower bounds by global optimization algorithm. Finally, through a numerical study, the importance p-boxes of various CCF events are calculated and the changing diagrams about importance are gotten, the results identify the effectiveness of the proposed method. This paper only measures the importance of various CCF events to system reliability, in the future work, the importance of coupling common causes to system reliability should be further investigated which can provide a guidance for system design and maintenance.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China under contract No. 51805073 and U1830207, the Chinese Universities Scientific Fund under contract No. ZYGX2018J061, the Sichuan Science and Technology Project under contract No. 2019JDJQ0015. Jinhua Mi wishes to acknowledge the financial support of the China Scholarship Council.

References

1. Alizadeh S, Sriramula S. Impact of common cause failure on reliability performance of redundant safety related systems subject to process demand. Reliability Engineering & System Safety, 2018; 172: 129-150, https://doi.org/10.1016/j.ress.2017.12.011.
2. Birnbaum ZW. On the importance of different components in a multicomponent system. In Multivariate Analysis, II, Academic Press, New York 1968: 581-592, https://doi.org/10.21236/AD0670563.
3. Boudali H, Dugan JB. A continuous-time Bayesian network reliability modeling, and analysis framework. IEEE Transactions on Reliability

2006; 51(1): 86-97, https://doi.org/10.1109/TR.2005.859228.
4. Dannert MM, Fau A, Fleury RMN, Broggi M. A probability-box approach on uncertain correlation lengths by stochastic finite element method. Proceeding in Applied Mathematics and Mechanics 2018; 18(1): e201800114, https://doi.org/10.1002/pamm.201800114.
5. Feng G, George-Williams H, Patelli E, Coolen FPA, Beer M. An efficient reliability analysis on complex non-repairable systems with common-cause failures. Safety and Reliability-Safe Societies in a Changing World. CRC Press 2018: 2531-2537, https://doi.org/10.1201/9781351174664-318.
6. Ferson S, Hajagos J, Berleant D, Zhang J, Tucker WT, Ginzburg L, Oberkampf W. Dependence in Dempster-Shafer theory and probability bounds analysis. Sandia National Laboratories 2004.
7. Ferson S, Nelsen RB, Hajagos J, Berleant DJ, Zhang J, Tucker WT, Oberkampf WL. Dependence in probabilistic modeling, Dempster-Shafer theory, and probability bounds analysis. Sandia National Laboratories 2015.
8. Karanki DR, Kushwaha HS, Verma AK, Ajit S. Uncertainty analysis based on probability bounds (p-box) approach in probabilistic safety assessment. Risk Analysis: An International Journal 2009; 29(5): 662-675, https://doi.org/10.1111/j.1539-6924.2009.01221.x.
9. Le Duy TD, Vasseur D. A practical methodology for modeling and estimation of common cause failure parameters in multi-unit nuclear PSA model. Reliability Engineering & System Safety 2018; 170: 159-174, https://doi.org/10.1016/j.ress.2017.10.018.
10. Li H, Huang HZ, Li YF, Zhou J, Mi J. Physics of failure-based reliability prediction of turbine blades using multi-source information fusion. Applied Soft Computing 2018; 72: 624-635, https://doi.org/10.1016/j.asoc.2018.05.015.
11. Li YF, Huang HZ, Mi J, Peng W, Han X. Reliability analysis of multi-state systems with common cause failures based on Bayesian network and fuzzy probability. Annals of Operations Research 2019: 1-15, https://doi.org/10.1007/s10479-019-03247-6.
12. Li YF, Mi J, Liu Y, Yang YJ, Huang HZ. Dynamic fault tree analysis based on continuous-time Bayesian networks under fuzzy numbers. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2015; 229(6): 530-541, https://doi.org/10.1177/1748006X15588446.
13. Li YF, Mi J, Huang HZ, Zhu SP, Xiao N. Fault tree analysis of train rear-end collision accident considering common cause failure. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(4): 403-408.
14. Mi J, Cheng Y, Song Y, Bai L, Chen K. Application of dynamic evidential networks in reliability analysis of complex systems with epistemic uncertainty and multiple life distributions. Annals of Operations Research 2019: 1-23, https://doi.org/10.1007/s10479-019-03211-4.
15. Mi J, Li YF, Liu Y, Yang YJ, Huang HZ. Belief universal generating function analysis of multi-state systems under epistemic uncertainty and common cause failures. IEEE Transactions on Reliability 2015; 64(4): 1300-1309, https://doi.org/10.1109/TR.2015.2419620.
16. Mi J, Li YF, Peng W, Yang YJ, Huang HZ. Fault tree analysis of feeding control system for computer numerical control heavy-duty horizontal lathes with multiple common cause failure groups. Journal of Shanghai Jiaotong University (Science) 2016; 21(4): 504-508, https://doi.org/10.1007/s12204-016-1755-7.
17. Mi J, Li YF, Peng W, Huang HZ. Reliability analysis of complex multi-state system with common cause failure based on evidential networks. Reliability Engineering & System Safety 2018; 174: 71-81, https://doi.org/10.1016/j.ress.2018.02.021.
18. Mi J, Li YF, Yang YJ, Peng W, Huang HZ. Reliability assessment of complex electromechanical systems under epistemic uncertainty. Reliability Engineering & System Safety 2016; 152: 1-15, https://doi.org/10.1016/j.ress.2016.02.003.
19. O'Connor AN. A general cause based methodology for analysis of dependent failures in system risk and reliability assessments, University of Maryland, 2013.
20. O'Connor AN, Mosleh A. Extending the alpha factor model for cause based treatment of common cause failure events in PRA and event assessment. Proceedings of the 12th probabilistic safety assessment & management conference (PSAM12) 2014.
21. Qiu S, Ming HXG, Hou Y. An evidential network-based method for common-cause failure analysis under uncertainty. In Safety and Reliablity-Safe Societies in a Changing World, Haugen et al. (eds), 2018: 2365-2372, https://doi.org/10.1201/9781351174664-297.
22. Sakurahara T, Schumock G, Reihani S, Kee E, Mohaghegh Z. Simulation-informed probabilistic methodology for common cause failure analysis. Reliability Engineering & System Safety 2019; 185: 84-99, https://doi.org/10.1016/j.ress.2018.12.007.
23. Song Y, Mi J, Cheng Y, Bai L, Wang X. Application of discrete-time Bayesian network on reliability analysis of uncertain system with common cause failure. Quality and Reliability Engineering International 2019; 35(4): 1025-1045, https://doi.org/10.1002/qre.2443.
24. Tucker WT, Ferson S. Probability bounds analysis in environmental risk assessment. Applied Biomathematics, Setauket, New York, 2003.
25. Wang C, Matthies HG, Xu M, Li Y. Dual interval-and-fuzzy analysis method for temperature prediction with hybrid epistemic uncertainties via polynomial chaos expansion. Computer Methods in Applied Mechanics and Engineering 2018; 336: 171-186, https://doi.org/10.1016/j.cma.2018.03.013.
26. Wang C, Xing L, Levitin G. Explicit and implicit methods for probabilistic common-cause failure analysis. Reliability Engineering & System Safety 2014; 131: 175-184, https://doi.org/10.1016/j.ress.2014.06.024.
27. Wang C, Xing L, Levitin G. Probabilistic common cause failures in phased-mission systems. Reliability Engineering & System Safety 2015; 144: 53-60, https://doi.org/10.1016/j.ress.2015.07.004.
28. Wang L, Xiong C, Yang Y. A novel methodology of reliability-based multidisciplinary design optimization under hybrid interval and fuzzy uncertainties. Computer Methods in Applied Mechanics and Engineering 2018; 337: 439-457, https://doi.org/10.1016/j.cma.2018.04.003.
29. Wei P, Lu Z, Song J. Variable importance analysis: a comprehensive review. Reliability Engineering & System Safety 2015; 142: 399-432, https://doi.org/10.1016/j.ress.2015.05.018.
30. Xie L, Lundteigen MA, Liu YL. Safety barriers against common cause failure and cascading failure: literature reviews and modeling strategies, 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE) 2018: 122-127, https://doi.org/10.1109/IEEM.2018.8607769.
31. Zhang M, Zhang Z, Mosleh A, Chen S. Common cause failure model updating for risk monitoring in nuclear power plants based on alpha factor model. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2017; 231(3): 209-220, https://doi.org/10.1177/1748006X16689542.
32. Zhang X, Gao H, Huang HZ, Li YF, Mi J. Dynamic reliability modeling for system analysis under complex load. Reliability Engineering & System Safety, 201;180: 345-351, https://doi.org/10.1016/j.ress.2018.07.025.
33. Zhu P, Han J, Liu L, Lombardi F. A stochastic approach for the analysis of dynamic fault trees with spare gates under probabilistic common

cause failures. IEEE Transactions on Reliability 2015; 64(3): 878-892, https://doi.org/10.1109/TR.2015.2419214.
34. Zuo L, Xiahou T, Liu Y. Reliability assessment of systems subject to interval-valued probabilistic common cause failure by evidential networks. Journal of Intelligent & Fuzzy Systems 2019; 36(4): 3711-3723, https://doi.org/10.3233/JIFS-18290.

Jinhua Mi
School of Automation Engineering,
University of Electronic Science and Technology of China,
No.2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu, Sichuan, 611731, P.R. China

Yan-Feng Li
Center for System Reliability and Safety,
University of Electronic Science and Technology of China,
No.2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu, Sichuan, 611731, P.R. China

Michael Beer
Matteo Broggi
Institute for Risk and Reliability,
Leibniz University Hannover,
Callinstr. 34, Hannover, 30167, Germany

Yuhua Cheng
School of Automation Engineering,
University of Electronic Science and Technology of China,
No.2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu, Sichuan, 611731, P.R. China

Emails: jinhuami@uestc.edu.cn, yanfengli@uestc.edu.cn, beer@irz.uni-hannover.de, broggi@irz.uni-hannover.de, yhcheng@uestc.edu.cn,

Article citation info:
JÓŹWIAK A. OWCZAREK P, PROCHOWSKI L, ŚWIDERSKI A. Analysis of the impact of the use time of n1 motor vehicles on the economic efficiency of their maintenance. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 121–129, http://dx.doi.org/10.17531/ein.2020.1.14.

Arkadiusz JÓŹWIAK
Paulina OWCZAREK
Leon PROCHOWSKI
Andrzej ŚWIDERSKI

ANALYSIS OF THE IMPACT OF THE USE TIME OF N1 MOTOR VEHICLES ON THE ECONOMIC EFFICIENCY OF THEIR MAINTENANCE

BADANIE WPŁYWU CZASU WYKORZYSTANIA SAMOCHODÓW KATEGORII N1 NA EFEKTYWNOŚĆ EKONOMICZNĄ ICH EKSPLOATACJI*

The efficiency of operation of motor vehicles with a DMC (Permissible Laden Mass) <3.5 tonnes is considered. These are vehicles belonging motor vehicles of category N1, usually referred to as delivery vehicles. The results of observations on the implementation of transport orders in 7 transport companies from the MŚP (Small and Middle-size Companies) sector were used to conduct the effectiveness analysis. The research group covered 24 vehicles that implementation transport orders in the urban zone and in the immediate vicinity of the city. Information was collected on a monthly basis.During the analysis of economic efficiency the income measures (absolute and relative) were used. The calculations were carried out using the model of the vehicle operation process in the form of a neural network, in which a set of 12 input variables and 3 output variables were taken into account. Using the Statistica 13.3 computer program and defining the group and factors describing the process of implementation of individual transport tasks, the developed neural network model enabled searching for the impact of selected operational factors on the economic efficiency of N1 category cars.The calculations showed a significant impact of the number of vehicle days in a month, the weight of the load, as well as the time of year. The obtained calculation results showed the specific features of the impact of the number of working days on revenue in a transport company. The increase in the number of working days favors the increase in income in a limited way, and this restriction depends, among others since the time of year.

***Keywords**: operation of vehicles, motor vehicles of category N1, economic efficiency, neural networks.*

Rozważa się efektywność eksploatacji samochodów ciężarowych o DMC < 3,5 tony. Są to pojazdy należące do kategorii N1 (według Dyrektywy 2007/46/WE) zwykle nazywane samochodami dostawczymi. Do prowadzonej analizy efektywności wykorzystano wyniki obserwacji z realizacji zleceń przewozowych w 7 firmach transportowych z sektora MŚP. Grupa badawcza objęła 24 pojazdy, które wykonywały zadania transportowe w strefie miejskiej i w najbliższym otoczeniu miasta. Informacje gromadzono w cyklach miesięcznych. Podczas analizy efektywności ekonomicznej zastosowano kilka miar przychodu (bezwzględny i względny). Obliczenia prowadzono przy wykorzystaniu modelu procesu eksploatacji pojazdów w postaci sieci neuronowej, w której brano pod uwagę zbiór 12 zmiennych wejściowych i 3 zmienne wyjściowe. Stosując program komputerowy Statistica 13.3 oraz zdefiniowanie grupy i czynniki opisujące proces realizacji poszczególnych zadań transportowych, opracowany model sieci neuronowej umożliwił poszukiwanie wpływu wybranych czynników eksploatacyjnych na efektywność ekonomiczną samochodów kategorii N1. Przeprowadzone obliczenia pokazały istotny wpływ liczby dni pracy pojazdów w miesiącu, masę ładunku, a także porę roku. Uzyskane wyniki obliczeń pokazały specyficzne cechy wpływu liczby dni pracy na przychód w firmie transportowej. Wzrost liczby dni pracy sprzyja wzrostowi przychodu w sposób ograniczony, a to ograniczenie zależy m.in. od pory roku.

***Słowa kluczowe**: eksploatacja samochodów, pojazdy samochodowe kategorii N1, efektywność ekonomiczna, sieci neuronowe.*

1. Introduction

The specificity of the use of N1 motor vehicles constantly raises a lot of controversy, and in the area of legal conditions there are still a number of ambiguities, which affects the existence of rather negligible amount of literature on this subject. Every month there is new information regarding the statistics of the Main Road Transport Inspectorate about the results of N1 category motor vehicles inspections. For several years, the percentage of vehicles with a load exceeding their permissible capacity in relation to the category N1 vehicles inspected has remained on average at the 93% level (table 1) [19]. It is a fact that the Road Transport Inspectorate usually checks heavy goods vehicles over 3.5 tons, while N1 category vehicles only when it considers that there is a clear suspicion of committing a specific offense. The number of vehicles inspected is negligible, but the percentage ratio of vehicles with a load exceeding their permissible load capacity of up to 3.5 tons may indicate the existence of a complex problem that should be subject to detailed analysis.

Therefore, it becomes justified to be interested in the subject of increasing the profits from operation of the N1 category motor vehicles. Reproducing and simplifying real phenomena in the form of a model becomes an important element in the search for effective methods to

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

Table 1. Percentage share of vehicles with a load exceeding their permissible capacity in relation to the N1 category vehicles tested [19]

	2014	2015	2016	2017
Number of all vehicles with GVW up to 3.5 t	2 399 323	2 447 764	2 515 751	2 574 312
Number of vehicles checked with GVW up to 3.5 t	6 685	6 507	7 594	10 048
Number of tickets issued for vehicles with a load exceeding their maximum capacity	6 135	6 172	6 172	9 396
Percentage share	92%	95%	92%	94%

describe problems and disturbances in the process of operating and servicing motor vehicles. As a result, it makes it easier to find ways to increase the economic efficiency of transport companies.

In the article, the subject of research is the effectiveness of operation of the N1 category cars, which according to [27] are defined as vehicles designed and constructed for transporting loads and having a maximum total weight not exceeding 3.5 tons (GVW up to 3.5 t). In practice, this type of vehicle is referred to as delivery vehicles, which is why this term will be used interchangeably later in the article.

Planning and implementing the process of using motor vehicles in a complex transport system is associated with solving multi-criteria decision problems, which focus on, among the others minimizing costs and achieving maximum profit. This issue concerns the issues of two basic elements of the operation process, i.e. the use and maintenance of motor vehicles. Effective use of transport means in every enterprise is one of the main ways to achieve competitive advantage.

Extensive analyzes of the motor vehicles operation process most often relate to minimizing the costs associated with their use and ensuring maximum reliability of the transport system, as well as the impact of operating the vehicles on the natural environment [21], or safety aspect [20]. Whereas the assessment of the intensity of motor vehicle operation is carried out taking into account, inter alia, mileage values, engine capacity, vehicle's age [18], repair costs, revenues [16], technical availability, intensity of motor vehicle use [6]. Due to the random nature of vehicle failures, the knowledge of stochastic processes is necessary to maintain their efficient and safe operation [2].

Modelling and assessment of such complex processes based on classic mathematical models and techniques of reliability theory may be difficult to carry out and not bring the expected results due to the large amount of quantitative and qualitative data and due to the dynamically changing conditions of the vehicle operation system. In such a case, other computational methods are proposed, e.g. models using Markov processes or reliable phase diagrams, a Petri network model or Monte Carlo simulation processes [13], an algorithm of resistance clonal selection [5].

Considering the degree of complexity of the studied problem and the increasing use of artificial intelligence methods to solve this type of task, the goal of this study is to assess the economic efficiency of transport companies using a neural network. The evaluation is focused on transport companies operating for package cargo transport in urban and suburban areas. The work uses several measures of revenue in transport activities described later in this work.

Wherever there are no grounds for linear approximation of occurring phenomena and processes, usually when solving difficult and troublesome evaluation issues, including efficiency of car operation, it may be rational to refer to neural networks or other artificial intelligence algorithms (i.e. models that map non-linear relationships) [7], [9], [10], [24]. Artificial neural networks are one of the techniques used by artificial intelligence. There are also other uses of artificial intelligence in transport. For example: for assessing the quality of transport means, for optimizing travel routes [11], or for evaluating the configuration of transport service management [23].

2. Cost analysis of a transport enterprise with a fleet of N1 category motor vehicles

The transport service provider is still selected by the minimum price criterion as the first choice. High competition and constantly growing requirements of the transport market force carriers to constantly search for methods to minimize the costs of a transport company. Generating revenues at the transport companies is primarily based on the intensity of the vehicle operation. As a rule, they are proportional to the number of kilometers traveled, load weight or hours worked. The intensive operation of vehicles generates not only an increase in revenues, but also costs, which is why it is extremely important to carry out their detailed analysis.

In enterprises, including transport ones, one of the most commonly used cost sharing methods is their generic system, containing 7 groups, which are also the names of synthetic accounts: depreciation, consumption of materials and energy, external services, taxes and fees, salaries, social insurance and other benefits, other costs (generic ones).

According to many publications, the costs of external services represent the largest percentage in relation to all costs of the transport company [15]. In this study, the research subjects are micro, small and medium enterprises (micro and SMEs), therefore the cost structure will be slightly different from the general classification of generic type of costs of the enterprises. The reason for this could be, for example, the fact that micro and SMEs only have their own, not leased rolling stock, which provides transport services directly without the participation of outsourcing companies. Issues related in detail to the costs of road freight transport enterprises are of interest to many authors [1], [4], [8], [12], [15], [26] who most often reduce them to four basic generic groups and determine their percentage values in relation to other costs:

- depreciation of 6% - 12%,
- operation 20% - 68%,
- drivers' remuneration 14% - 45%,
- remaining costs 12% - 30%.

Based on the cited analysis of the literature, the fig. 1 presents the shares of the three basic cost groups of automotive freight transport enterprises from the perspective of several authors.

For the purpose of achieving the research goal, the data was collected, which was classified into four groups of factors: utility, season, service, economic.

3. Research method and object

Tasks carried out at the transport companies that provide services in Poland have been examined. The operation of rolling stock, belonging to 7 different transport companies from the SME sector, involves the implementation of transport tasks in accordance with the needs of customers. The research group includes 24 N1 category motor vehicles, 5 models: Renault Master, Renault Mascott, Citroen Jumper and Fiat Ducato. The study only took into account technical data that had an impact on the aforementioned factors.

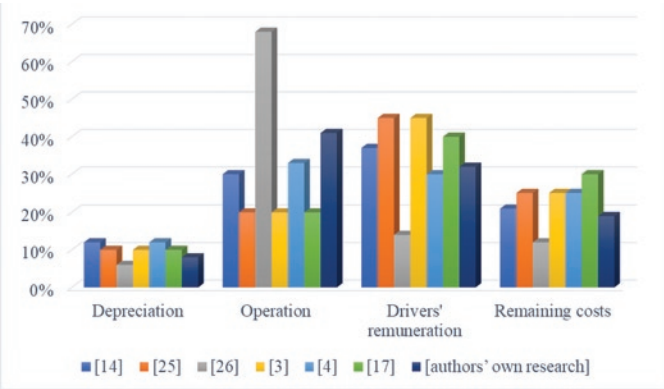


Fig. 1. Comparison of the share of selected cost groups of road freight transport enterprises

The following measures of economic efficiency have been defined:

- revenue $[Y_M^E]$ - expressed as the difference between the monthly value and the monthly operating costs,
- relative revenue $[Y_L^E]$ - expressed as the ratio of the monthly value of orders to the monthly mileage,
- relative profit $[Y_W^E]$ - expressed as the ratio of the revenue to the monthly mileage. There are 4 main groups of factors taken into account in the research:
- Y^U - a group of quantities describing the factors of operating motor vehicles: characterizing the manner and intensity of the work performed,
- Y^K - a group of quantities describing the seasons: defining the external conditions in which the vehicle is used,
- Y^O - a group of quantities describing the maintenance of motor vehicles: concerning the maintenance strategy and its effects,
- Y^E - a group of quantities describing economic factors: related to the costs and profitability of carrying out transport tasks.

A set of factors describing the above groups are shown in table 2.

For such defined groups, a real data from one year of vehicles operation (2017) was collected in monthly cycles. These vehicles droved mainly in urban traffic with a few routes outside the city, within this country. The data was obtained from transport orders in the period under review, analyses of the service expertises and interviews with experts (dispatchers, drivers, service technicians, mechanics). The 156 observations of the above factors were made for each vehicle. This way, 3744 data was collected, which was used to model the operation process using a neural network.

Table 2. Set of factors used to modelling maintenance process

	Designation of groups and factors		Units of measurement
	Y^U	**Group: factors of the motor vehicles operation**	
1	Y_D^U	number of days of the vehicle use in a month	number
2	Y_R^U	monthly vehicle's mileage	kilometres
3	Y_J^U	monthly vehicle's driving time	minutes
4	Y_C^U	monthly vehicle's working time	minutes
5	Y_P^U	average fuel consumption	litres/100 km
6	Y_M^U	average daily load weight	kilograms
7	Y_E^U	percentage value of the capacity utilization	%
	Y^K	**Group: time of year**	
8	Y_W^K	time of year	season 1, season 2, season 3
	Y^O	**Group: servicing activities of motor vehicles**	
9	Y_P^O	service fluid refilling	performed, not performed
10	Y_K^O	tire service	performed, not performed
11	Y_H^O	brakes service	performed, not performed
	Y^E	**Group: economic factors**	
12	Y_Z^E	monthly value of orders	PLN
13	Y_K^E	monthly operating cost	PLN
14	Y_M^E	month revenue from the implementation of transport services	PLN
15	Y_L^E	relative unit revenue	PLN/ km
16	Y_W^E	relative unit profit	PLN/ km

4. Neural modelling

When creating the neural network, some of the signals from the table 2 were used, these are:

- quantitative input ones: Y_D^U, Y_R^U, Y_J^U, Y_C^U, Y_P^U, Y_M^U, Y_E^U, Y_K^E,
- quantitative output: Y_M^E, Y_L^E, Y_W^E.

Use the results, among the others [22] of the scientific work, a Multilayer Perceptron and teaching algorithms were used: conjugate gradients; the fastest fall and BFGS (Broyden - Fletcher - Goldfarb - Shanno). The division of the data set into parts in neural modelling was adopted:

- 80% - teaching set used to modify weights,
- 10% - test set for ongoing monitoring of the teaching process,
- 10% - a validation set for assessing the quality of the network after the teaching process.

After determining the input signals, output signals and network parameters, the neural network teaching process was carried out using the Statistica 13.3 computer program. Examples of its results are presented in table 3.

Table 3. Sample results of the neural network teaching process

ID	Network name	Teaching quality	Testing quality	Validation quality	Teaching algorithm	Hidden activation	Output activation
1	MLP 17-11-3	0,796875	0,746498	0,874968	BFGS 28	Exponential	Exponential
2	MLP 17-16-3	0,785931	0,764052	0,874283	BFGS 50	Sinus	Logistic
3	MLP 17-36-3	0,775433	0,804533	0,828288	BFGS 17	Sinus	Exponential
4	MLP 17-30-3	0,762021	0,760954	0,861609	BFGS 23	Sinus	Linear
5	MLP 17-9-3	0,649014	0,612412	0,770544	BFGS 12	Sinus	Tanh
6	MLP 17-6-3	0,755626	0,766049	0,867752	BFGS 51	Linear	Tanh
7	MLP 17-8-3	0,767523	0,742646	0,876113	BFGS 66	Sinus	Tanh
8	MLP 17-31-3	0,776866	0,784677	0,872670	BFGS 36	Linear	Logistic
9	MLP 17-26-3	0,730811	0,783034	0,825576	BFGS 10	Exponential	Sinus
10	MLP 17-9-3	0,807437	0,754784	0,865342	BFGS 38	Logistic	Linear
11	MLP 17-26-3	0,794570	0,754735	0,874439	BFGS 24	Tanh	Logistic
12	MLP 17-3-3	0,796772	0,816948	0,850614	BFGS 52	Logistic	Sinus
13	MLP 17-29-3	0,776284	0,783517	0,874145	BFGS 31	Linear	Logistic
14	MLP 17-19-3	0,813487	0,752890	0,844996	BFGS 45	Logistic	Sinus
15	MLP 17-23-3	0,776775	0,783067	0,873542	BFGS 32	Linear	Logistic
16	MLP 17-17-3	0,838238	0,761414	0,700263	BFGS 64	Tanh	Tanh
17	MLP 17-7-3	0,768890	0,799204	0,819687	BFGS 17	Linear	Exponential
18	MLP 17-6-3	0,780286	0,768095	0,886923	BFGS 28	Logistic	Linear

5. Validation of the neural network model and calculations results

The structure of the best neural network took the form of MLP 17-19-3, which means 17 neurons in the input layer, 19 neurons in the hidden layer and 3 neurons in the output layer (fig. 2).

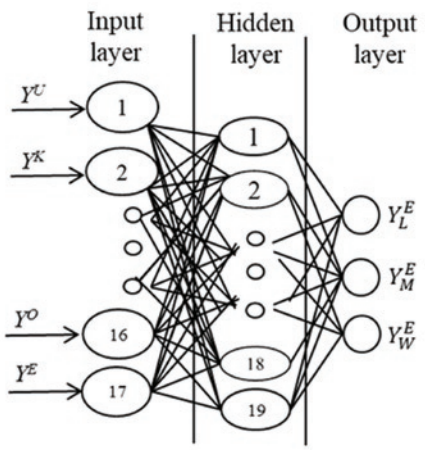


Fig. 2. Structure of the created MLP 17-19-3 network

Since among the input signals, the qualitative input signals have appeared, the total number of neurons at the input represent the sum of all quantitative and qualitative signals, broken down into their individual values. table 4 shows the input signals of the selected neural network.

In the table 3, the MLP 17-19-3 network teaching quality was estimated at around 81% probability of indicating the correct answer, i.e. the adopted measure of economic efficiency. Testing quality - at 75% level and validation quality - at 85%. The BFGS 45 algorithm turned

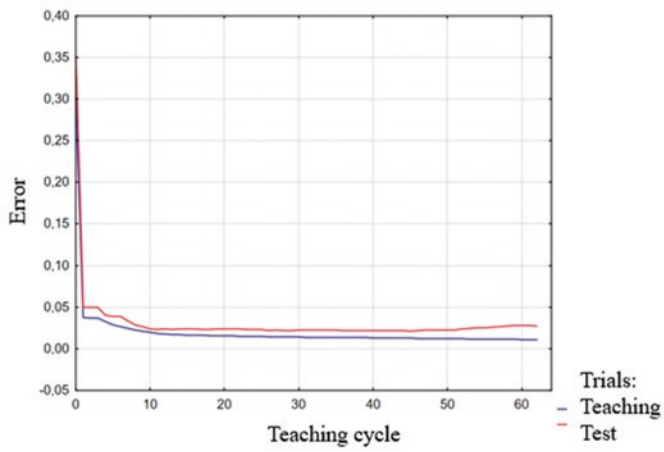


Fig. 3. Results of MLP 17-19-3 neural network teaching

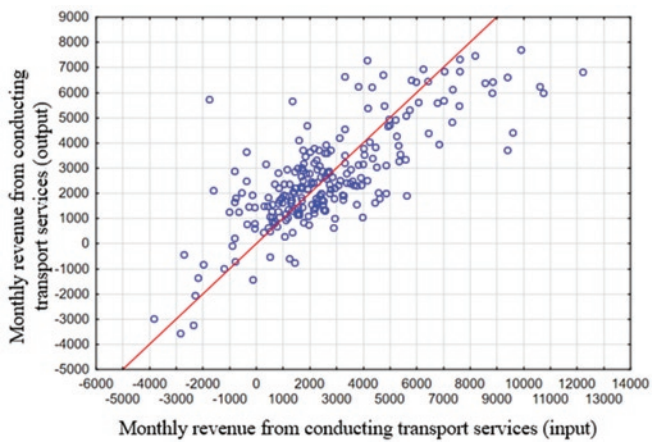


Fig. 4. Dispersion of the dependent variable of the MLP 17-19-3 neural network

Table 4. MLP 17-19-3 neural network individual input signals

	Designation of groups and factors	Units of measurement	ID of the neuron	Value of the neuron
1	Y_D^U	number	1	Y_D^U
2	Y_R^U	kilometers	2	Y_R^U
3	Y_J^U	minutes	3	Y_J^U
4	Y_C^U	minutes	4	Y_C^U
5	Y_P^U	litres/100 km	5	Y_P^U
6	Y_M^U	kilograms	6	Y_M^U
7	Y_E^U	%	7	Y_E^U
8	Y_W^K	season 1, season 2, season 3	8	$Y_W^K\ seazon1$
			9	$Y_W^K\ seazon2$
			10	$Y_W^K\ sezaon3$
9	Y_P^O	completed, not completed	11	$Y_P^O\ completed$
			12	$Y_P^O\ not\,completed$
10	Y_K^O	completed, not completed	13	$Y_K^O\ completed$
			14	$Y_K^O\ not\,completed$
11	Y_H^O	completed, not completed	15	$Y_H^O\ completed$
			16	$Y_H^O\ not\,completed$
12	Y_K^E	PLN	17	Y_K^E

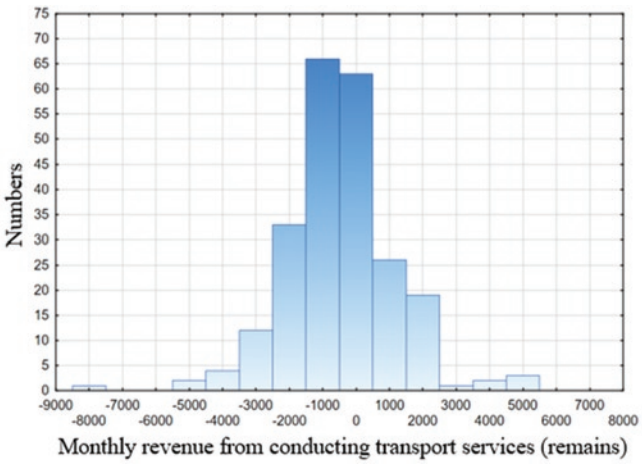


Fig. 5. Distribution of the residues of MLP 17-19-3 neural network

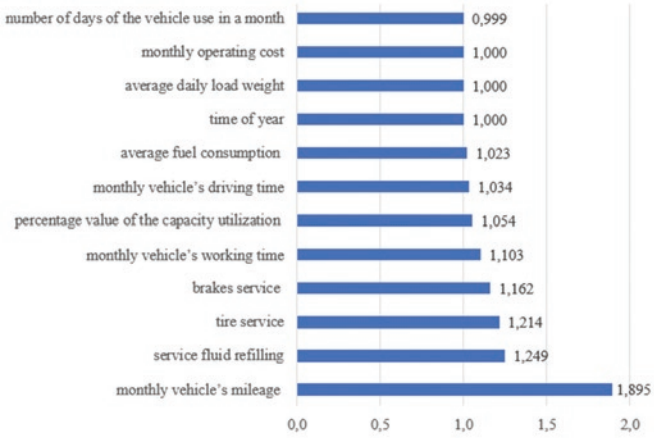


Fig. 6. Global sensitivity analysis for the MLP 17-19-3 neural network

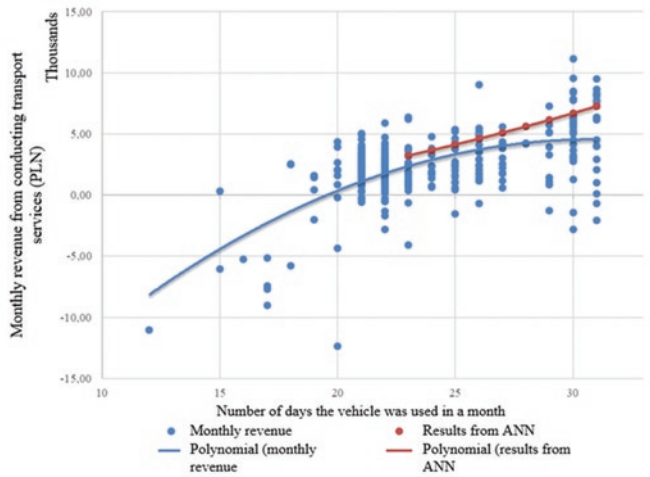


Fig. 7. Monthly revenue in relation to the number of days of the vehicle's operation

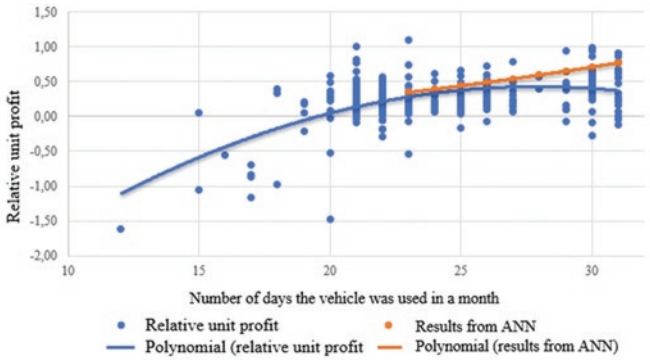


Fig. 9. Relative unit profit in respect to the number of days of operation of the vehicle

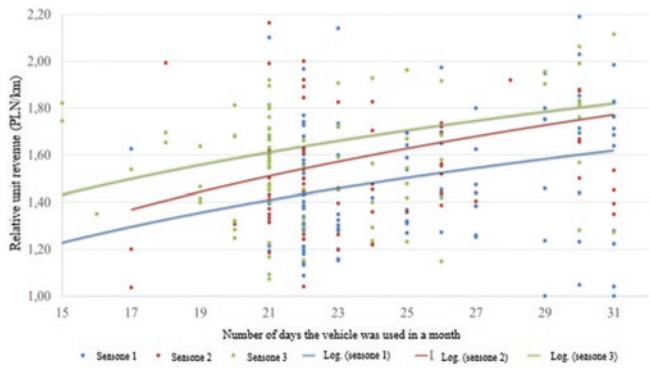


Fig. 11. Relative unit revenue in relation to the number of days of operation of the vehicle in respect to Y^K

out to be the best teaching algorithm. The positive result of teaching the neural network is demonstrated, by the teaching graph (fig. 3). It shows that the best network structure was found in the 43-rd cycle; the share of incorrect answers was 19% and the error was estimated at 0.002. Also the course of the dispersion changes, shown in fig. 4, indicates a positive outcome of teaching the network.

The results of the calculations presented in fig. 4 show the dispersion between the forecasted revenue value (the result of the calculation in the network during teaching) and its actual value. The visible gathering of the dispersion value around zero is a good result of the model calculations.

The histogram presented in fig. 5 (distribution of residues, i.e. differences between the output variable and its prediction) shows

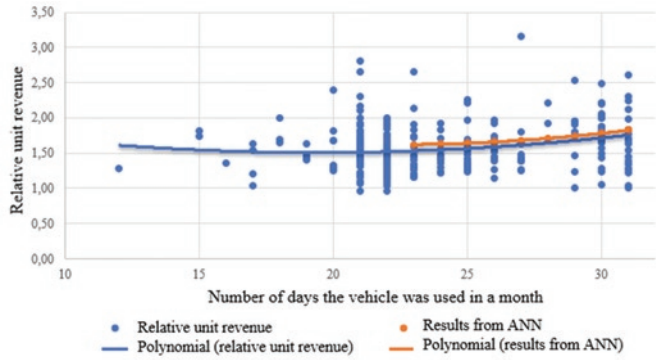


Fig. 8. Relative unit revenue in relation to the number of days of the vehicle's operation

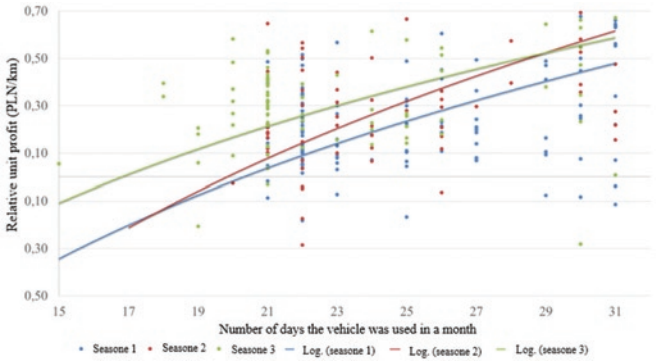


Fig. 10. Relative unit profit in relation to the number of days of operation of the vehicle in respect to Y^K

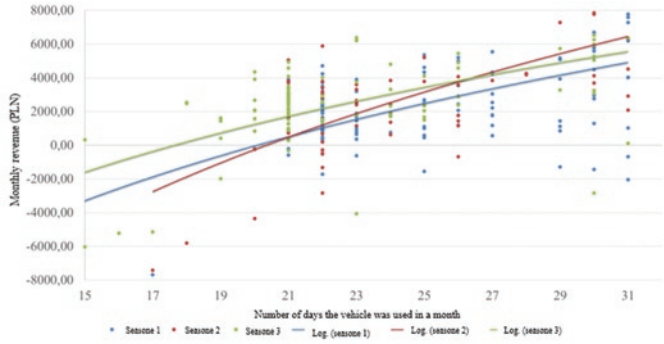


Fig. 12. Monthly revenue in relation to the number of days of operation of the vehicle in respect to Y^K

the number of results of the scatter calculation near zero, which also means a high level of reproducing of the output signals.

In the next stage, a sensitivity analysis was carried out, which involved checking how the network error behaves when input signals are modified. In this calculation, the input signal values are replaced by the average of this signal from the teaching set. After inputting such modified data, the network error was checked. If the error has increased significantly, it means that the network is very sensitive to the signal.

The global sensitivity analysis reflects the impact of individual network input variables on the output signals (fig. 6). These calculations have shown that the greatest impact on the output signals of the neural network have: the monthly number of days of the vehicle's operation, the cost of operation, load weight and time of year.

Based on the selected neural network and the data collected for the neural network teaching process, the trends of changes in the value of efficiency measures are shown, namely: monthly revenue,

Table 5. Example results of the calculations of revenue, unit relative income and unit relative profit for a data set not used in the teaching process

P_Y^X	Y_D^U	Y_R^U	Y_J^U	Y_C^U	Y_P^U	Y_M^U	Y_E^U	Y_K^E	Y_W^K	Y_P^O	Y_K^O	Y_H^O	Y_M^E	Y_L^E	Y_W^E
1	30	11360	425	480	15	1661	1,33	12873,13	season 2	yes	yes	yes	5976,87	1,66	0,53
2	31	13331	516	553	15	1771	1,61	15669,06	season 2	yes	yes	yes	2920,94	1,39	0,22
3	24	10127	422	454	15	1408	1,14	13172,67	season 2	yes	yes	yes	1787,33	1,48	0,18
4	30	13750	516	581	15	1887	1,45	16232,06	season 2	yes	yes	yes	9537,95	1,87	0,69
5	30	9470	350	427	15	1697	1,31	14113,56	season 2	yes	yes	yes	3686,44	1,88	0,39
6	28	10527	376	442	15	1556	1,64	16041,24	season 2	yes	yes	yes	4178,76	1,92	0,40
7	17	10700	684	784	15	1588	1,22	20263,87	season 2	yes	yes	yes	-7413,87	1,20	-0,69
8	31	16362	528	602	15	1765	1,36	19235,94	season 2	yes	yes	yes	4514,06	1,45	0,28
9	17	10700	684	784	15	1349	1,23	20120,94	season 2	yes	yes	yes	-9020,94	1,04	-0,84
10	21	7824	447	484	13	743	0,99	9727,43	season 2	yes	yes	yes	722,57	1,34	0,09
11	21	9632	459	491	14	1114	1,11	11636,10	season 2	yes	yes	yes	1593,90	1,37	0,17
12	23	10560	515	635	14	1190	1,04	11430,36	season 2	yes	yes	yes	3319,64	1,40	0,31
13	22	6990	350	430	14	1162	1,01	9105,71	season 2	yes	yes	yes	3794,29	1,85	0,54
14	21	8297	395	459	14	1263	1,05	10928,23	season 2	yes	yes	yes	1541,78	1,50	0,19
15	21	9090	433	497	14	1036	1,04	10968,52	season 2	yes	yes	yes	1881,48	1,41	0,21
16	21	8000	419	520	14	771	1,03	9847,90	season 2	yes	yes	yes	952,10	1,35	0,12
17	31	11760	380	445	15	1454	1,32	15814,58	season 1	yes	yes	yes	4005,42	1,69	0,34
18	26	8600	331	378	15	1561	1,25	11779,84	season 1	yes	yes	no	5200,16	1,97	0,60
19	24	10570	441	538	15	1621	1,32	14242,71	season 1	yes	yes	yes	757,29	1,42	0,07
20	25	10980	440	503	15	1686	1,37	13236,73	season 1	yes	yes	no	5363,27	1,69	0,49
21	31	13230	471	552	15	1802	1,39	15843,46	season 1	yes	yes	yes	8356,54	1,83	0,63
22	30	11520	384	472	15	1620	1,25	16784,69	season 1	yes	yes	yes	2965,31	1,71	0,26
23	31	14500	468	493	15	1642	1,26	16733,81	season 1	yes	yes	yes	1016,19	1,22	0,07
24	31	14400	465	502	15	1513	1,16	15500,69	season 1	yes	yes	yes	8099,31	1,64	0,56
25	22	11290	514	571	15	1563	1,20	12614,45	season 1	yes	yes	yes	865,55	1,19	0,08
26	30	14930	498	603	15	1645	1,73	16990,02	season 1	yes	yes	yes	4509,98	1,44	0,30
27	31	17810	575	613	15	1558	1,64	19222,32	season 1	yes	yes	yes	-672,32	1,04	-0,04
28	30	12860	429	530	15	1645	1,73	17773,67	season 1	yes	yes	yes	5766,33	1,83	0,45
29	30	17040	569	674	15	1807	1,39	19713,65	season 1	yes	yes	yes	1286,35	1,23	0,08
30	12	6850	571	663	15	1779	1,37	19805,28	season 1	yes	yes	yes	-11055,28	1,28	-1,61
31	31	17980	580	618	15	1584	1,22	20049,37	season 1	yes	yes	yes	-2049,37	1,00	-0,11
32	25	14400	576	680	15	1596	1,45	14674,60	season 1	yes	yes	yes	4325,40	1,32	0,30
33	31	10936	423	488	15	1597	1,42	15795,00	season 3	yes	yes	yes	5475,83	2,05	0,50
34	30	10304	405	466	15	1588	1,38	14079,57	season 3	yes	yes	yes	4563,86	1,77	0,44
35	29	8372	358	413	15	1679	1,41	12553,19	season 3	no	no	no	3348,62	1,94	0,40
36	26	11752	503	598	15	1536	1,34	13012,55	season 3	yes	yes	yes	2188,21	1,33	0,19
37	26	10224	430	531	15	1430	1,30	13012,55	season 3	yes	yes	yes	3282,31	1,53	0,32
38	25	11422	472	548	15	1609	1,39	15104,38	season 3	yes	yes	yes	2220,28	1,56	0,19
39	25	8985	389	486	15	1473	1,34	12577,23	season 3	yes	yes	yes	1037,22	1,54	0,12
40	24	9015	391	488	15	1476	1,34	12564,77	season 3	no	no	no	1176,01	1,54	0,13
41	23	10850	435	500	15	1614	1,39	14725,20	season 3	no	no	no	3184,18	1,68	0,29
42	22	8902	441	524	14	1101	1,08	10680,49	season 3	no	no	no	1885,16	1,45	0,21
43	21	8784	435	523	14	1112	1,09	10635,87	season 3	no	no	no	2458,52	1,46	0,28
44	20	8761	434	523	14	1112	1,09	10628,81	season 3	no	no	no	1839,10	1,47	0,21
45	23	11580	475	547	15	1646	1,40	15622,78	season 3	no	no	no	2371,98	1,57	0,20
46	21	9126	434	518	14	1158	1,15	11411,48	season 3	no	no	no	1562,54	1,46	0,17
47	22	8903	437	530	14	1066	1,10	10670,73	season 3	no	no	no	1702,55	1,43	0,19
48	24	9126	409	502	15	1461	1,24	12763,30	season 3	yes	yes	yes	801,64	1,47	0,09

relative unit revenue and relative unit profit in respect to the number of days of vehicle operation (fig. 7, fig. 8, fig. 9).

The analysis conducted also shows that the selected network correctly reproduces the selected measures of economic efficiency. In order to carry out a detailed analysis of the impact of this factor on the output signal, the impact of the time of year on the values of efficiency measures in relation to the number of days of the vehicle operation was extracted (fig. 10, fig. 11, fig. 12).

Based on the interview with vehicles' drivers, the time of year was determined as the general conditions for meeting the orders, as well as driving comfort and safety:

- season 1 is assigned to months: May, June, July, August,
- season 2: March, April, September, October,

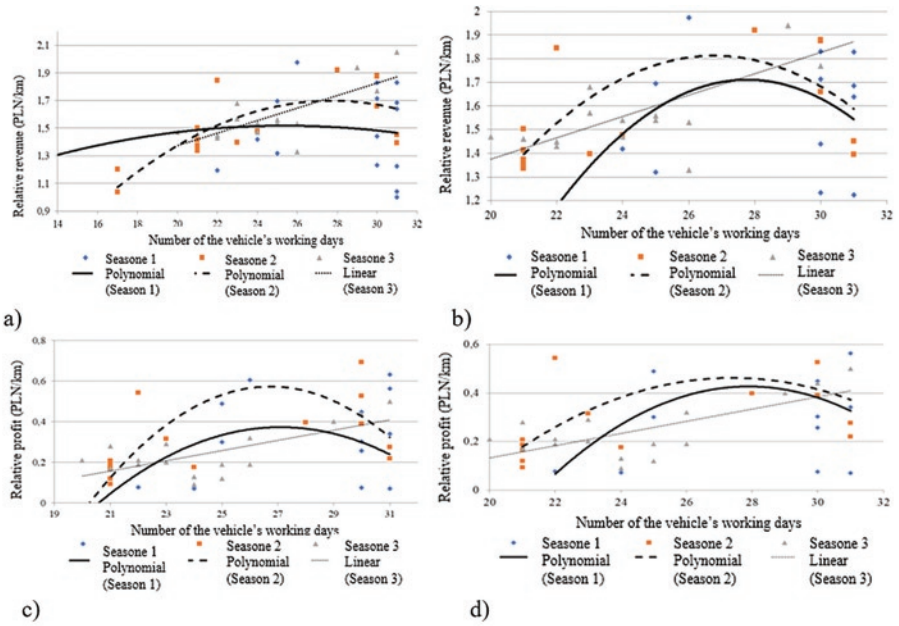


Fig. 13. Results of calculations the revenue and relative profit as a function of the number of days of the vehicle's work

• season 3: January, February, November, December.

The research showed that the highest values of efficiency measures are achieved during the operation of vans and fulfilling orders during season 3, and the lowest during season 1. This is confirmed by the fact of seasonality in providing transport services. Bad weather conditions determining the winter season is the time of increased use of vehicles due to the growing demand for services (season 3), with reduced supply of them. Good weather conditions are observed in the summer season, but then the demand for transport services decreases (season 1). The supply of services in the analyzed sector in this period is higher than the demand for services.

The verification of the proposed method was carried out based on the results of subsequent calculations made after inputting in the neural network a data not used in the teaching process. Examples of final results are presented in table 5.

Based on the global sensitivity analysis and the results obtained (table 5), it can be concluded that revenue increases with the increase in the number of days the vehicle works. Relative revenue and relative profit are calculated below in two ways: taking into account all variants of orders together with those generating a loss and taking into account only variants generating profit from order execution.

The studies have also shown that the season is also an important factor influencing the relationship between the number of working days and revenue or profit. The season determines the rate of increase in relative income and relative profit. This rate is high in the range of 10-22 days of vehicle work and becomes moderate in the range of 23-25 days of work per month. However, increasing the number of working days above 26-27 no longer results in an increase in benefits. The results presented in fig. 13c) confirm that in order to achieve relative profit vehicles should be used not less than 20 days a month, while the analysis of predictions without taking into account orders bringing losses (fig. 13d) confirms that it is over 21 days of work for season 2 and 22 days for season 1.

6. Summary

The results obtained and presented in the article allowed the statements that the adopted measures of economic efficiency have illustrated the impact of the number of vehicles' working days on the revenue and profit from transport services, and that the developed model is useful for predicting monthly revenue from transport services.

The results of the calculations provide the basis for the statement that increasing the number of days of the vehicle's work has a limited impact on the revenue growth process in the company. It is observed that positive income values are achieved with the number of working days over 19-20.

Both the number and type of data used in the neural network allowed to achieve high analysis results at the level of 80-90% efficiency.

The calculations results obtained results showed the specific features of the impact of the number of working days on the revenue in a transport company. The increase in the number of working days is conducive to increased revenue in a limited way, and this restriction depends on the season of the year.

The neural network model developed supports decision making in the implementation of transport processes, taking into account the economic efficiency of the motor vehicle operation process. Thus, the obtained results showed the usefulness of the adopted measures of economic efficiency and the model built to predict the economic results of the company's transport activities.

References

1. Aleksandrowicz P, Żółtowski B. Vehicle repair costs calculation systems. Polish Association for Knowledge Management. Warszawa: BEL Studio sp. z o.o., 2010.
2. Andrzejczak K, Młyńczak M, Selech J. Poisson-distributed failures in the predicting of the cost of corrective maintenance. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (4): 602-609, https://doi.org/10.17531/ein.2018.4.11.
3. Biesok G. Logistyka Usług. Warszawa: CeDeWu, 2013.
4. Bronk H. Cechy i układ kosztów w transporcie umożliwiające podejmowanie decyzji. Koszty i ceny w transporcie. Pomiar analiza. Szczecin: Zeszyty naukowe 813, 2014; 21 - 38.
5. Chen X, Xiao L, Zhang X, Xiao W, Li J. An integrated model of production scheduling and maintenance planning under imperfect preventive maintenance. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17 (1): 70-79, https://doi.org/10.17531/ein.2015.1.10.
6. Chłopek Z, Bebkiewicz K. Model of the structure of motor vehicles for the criterion of the technical level on account of pollutant emission. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (4): 501-507, https://doi.org/10.17531/ein.2017.4.2.
7. Coupek D, Gulec A, Lechler A,Verl A. Selective rotor assembly using fuzzy logic in the production of electric drives. Conference on Intelligent Computation in Manufacturing Engineering, 2014, https://doi.org/10.1016/j.procir.2015.06.074.
8. Gohari A, Matori N A, Yusof K W, Toloue I, Sholagberu A T. The effect of fuel price increase on transport cost of container transport vehicles.

International Journal of GEOMATE 2018; 15: 174-181, https://doi.org/10.21660/2018.50.30814.
9. Himanen V, Nijkamp P, Reggiani A. Neural networks in transport applications. Ashgate 1998. Reissued 2018 by Routledge, https://doi.org/10.4324/9780429445286.
10. Jóźwiak A. Application of Kohonen's Network in Logistics. Gospodarka Materiałowa i Logistyka 2017; 5: 258-271.
11. Kijek M, Brzeziński M, Gontarczyk M, Rykała Ł, Zelkowski J. Fuzzy Modeling of Evaluation Logistic Systems. Transport Means 2017; 2: 377-382.
12. Kleiner F, Friedrich H E. Development of a Transport Application based Cost Modelfor the assessment of future commercial vehicle concepts. Geneva: European Battery, Hybrid and Fuel Cell Electric Vehicle Congress, 2017.
13. Kowalski M, Magott J, Nowakowski T, Werbińska-Wojciechowska T. Analysis of transportation system with the use of Petri nets. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2011; 1 (49): 48-62.
14. Koźlak A. Ekonomika transportu. Teoria i praktyka. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego, 2010.
15. Mendyk E. Ekonomika transportu. Poznań: Wyższa Szkoła Logistyki, 2009.
16. Niewczas A, Rymarz J, Debicka E. Stages of operating vehicles with respect to operational efficiency using city buses as an example. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (1): 21-27, https://doi.org/10.17531/ein.2019.1.3.
17. Oziemski S. Efektywność eksploatacji maszyn. Podstawy techniczno - ekonomiczne. Radom: Biblioteka problemów eksploatacji, 1999.
18. Prochowski L. Evaluation of the process of mileage growth during the operation of motor trucks, in several categories of engine cubic capacity. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (3): 359-370, https://doi.org/10.17531/ein.2018.3.3.
19. Raporty z Internetu Głównego Inspektoratu Transportu Drogowego.
20. Rudyk T, Szczepański E, Jacyna M. Safety factor in the sustainable fleet management model. Archives of Transport 2019; 49: 103-114, https://doi.org/10.5604/01.3001.0013.2780.
21. Świderski A, Borucka A, Jacyna-Gołda I, Szczepański E. Wear of brake system components in various operating conditions of vehicle in the transport company. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (1): 1-9, https://doi.org/10.17531/ein.2019.1.1.
22. Świderski A, Jóźwiak A, Jachimowski R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (2): 292-299, https://doi.org/10.17531/ein.2018.2.16.
23. Świderski A. Studies and quality assurance neural modelling of the technical transport means. Archive of Transport. Polish Academy of Sciences Committee of Transport 2009; 21 (3-4).
24. Teodorovic D, Vukadinovic K. Traffic Control and Transport Planning:: A Fuzzy Sets and Neural Networks Approach. Springer Science+Business Media 2012.
25. Urbanyi-Popiołek I. Ekonomiczne i organizacyjne aspekty transportu. Bydgoszcz: Wyższa Szkoła Gospodarki, 2013.
26. Witkowski K, Tanona K. Analiza kosztów transportu drogowego. Logistyka 2013; 5: 411 - 416.
27. Zał. nr 2. do U. z dnia 20 czerwca 1997 r. Prawo o ruchu drogowym.

Arkadiusz JÓŹWIAK
Paulina OWCZAREK
Faculty of the Security, Logistics and Management
Military University of Technology
ul. gen. Sylwestra Kaliskiego 2, 00-908 Warsaw 46, Poland

Leon PROCHOWSKI
Faculty of Mechanical Engineering
Military University of Technology
ul. gen. Sylwestra Kaliskiego 2, 00-908 Warsaw 46, Poland

Andrzej ŚWIDERSKI
Motor Transport Institute
ul. Jagiellońska 80, 03-301 Warsaw, Poland

E-mails: arkadiusz.jozwiak@wat.edu.pl, paulina.owczarek@wat.edu.pl, leon.prochowski@wat.edu.pl, andrzej.swiderski@its.waw.pl

Article citation info:
PIELECHA I, PIELECHA J. Simulation analysis of electric vehicles energy consumption in driving tests. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 130–137, http://dx.doi.org/10.17531/ein.2020.1.15.

Ireneusz PIELECHA
Jacek PIELECHA

SIMULATION ANALYSIS OF ELECTRIC VEHICLES ENERGY CONSUMPTION IN DRIVING TESTS

SYMULACYJNA ANALIZA ENERGOCHŁONNOŚCI POJAZDÓW ELEKTRYCZNYCH W TESTACH BADAWCZYCH*

The assessment of energy flow through electric vehicle systems makes estimating their energy consumption possible. The article presents analyzes of the energy consumption of electric vehicles in selected driving tests (NEDC, WLTC and in real traffic conditions – RDC test) in relation to the vehicles different curb weight. The use of electric motors was also analyzed, providing their operating ranges, data of the energy flow in batteries and the change in their charge level. Simulation tests and analyzes were carried out using the AVL Cruise software. It was found that despite similar vehicle energy consumption values in NEDC and RDC testing, there are significant differences in energy flow in vehicle subsystems. The changes in the battery charge level per 100 km of test drive are similar in both the WLTC and RDC tests (6% difference); for the NEDC test, this difference is the greatest at 25% (compared to the previous tests). The energy consumption of electric vehicles depends significantly on the test itself; the values obtained in the tests are in the ranges of 10.1–13.5 kWh/100 km (NEDC test); 13–15 kWh/100 km (WLTC test) and 12.5–16.2 kWh/100 km in the RDC test. The energy consumption values in the NEDC and WLTC tests, compared to the RDC test, are approximately 20% and 10% lower, respectively. Increasing the vehicle mass increases the energy consumption (increasing the vehicle mass by 100 kg was found to increase the energy consumption by 0.34 kWh/100 km).

Keywords: *automotive vehicles, electric drive, high voltage batteries, energy flow, driving tests.*

Ocena przepływu energii przez układy pojazdów elektrycznych umożliwia oszacowanie ich energochłonności. W artykule przedstawiono analizy dotyczące zużycia energii pojazdów elektrycznych w wybranych testach jezdnych (NEDC, WLTC oraz w rzeczywistych warunkach ruchu – test RDC) w odniesieniu do zróżnicowanej masy pojazdów. Analizie poddano również wykorzystanie silników elektrycznych, przedstawiając mapy ich pracy, wielkości przepływu energii w akumulatorach oraz stopień zmiany ich naładowania. Badania i analizy symulacyjne wykonano z wykorzystaniem oprogramowania AVL Cruise. Stwierdzono, że mimo podobnych wartości energochłonności pojazdów w testach badawczych NEDC oraz RDC, to występują znaczące różnice przepływu energii w układach akumulacji pojazdów. Zmiany stopnia naładowania akumulatora odniesione do 100 km testu są zbliżone w testach WLTC oraz RDC (różnica 6%); dla testu NEDC różnica ta wynosi maksymalnie 25% (w odniesieniu do poprzednich testów). Energochłonność pojazdów elektrycznych jest silnie zależne od testu badawczego; wartości uzyskane w testach kształtują się na poziomie 10,1–13,5 kWh/100 km (test NEDC); 13–15 kWh/100 km (test WLTC) oraz 12,5–16,2 kWh/100 km w teście RDC. Wartości energochłonności w testach NEDC oraz WLTC są odpowiednio mniejsze o około 20% i 10% w odniesieniu do testu RDC. Zwiększenie masy pojazdu zwiększa zużycie energii (zwiększenie o 100 kg masy pojazdu zwiększa zużycie energii o 0,34 kWh/100 km).

Słowa kluczowe: *pojazdy samochodowe, napęd elektryczny, akumulatory wysokonapięciowe, przepływ energii, testy jezdne.*

1. Introduction

The push towards the reduction of fuel consumption of motor vehicles equipped with internal combustion engines (and hybrid drives) also drives an increase in the production of vehicles with electric propulsion systems, which leads to an increase in their share in the total number of vehicles in use. Changes to the requirements for reducing fuel consumption from the vehicle fleet by 2030 should be expected to contribute significantly to the development of electromobility [3] while also reducing greenhouse gas emissions.

Sales of electric vehicles (BEV – battery electric vehicle) on the global markets are increasing, however, their share in the overall car market is still not very large. The average European market share of EVs (electric vehicles) is about 2.5%, despite 200,000 units being sold in 2018 (Fig. 1). The most dynamic market is currently the Chinese market, with more than 800,000 new electric vehicle registrations in 2018. Despite this, the share of BEVs there is about 3.5%. In Europe, Norway is at the forefront with 45,000 new registrations in 2018 (the market share of BEVs there currently stands at 29%). In Poland, the share of BEVs is set at a low point of 0.4% [10]; at the end of July 2019, only 4009 electric vehicles were registered [14].

The last few years have seen a large increase in electric vehicle models in a whole range of segments (Fig. 2). The largest increase in models could be observed in the Chinese market – in all segments, where about 120 EV models are available. Compared to that there are only about 20 models on the European market – and half that on the American market [5]. Of the total vehicle sales in 2018 on the Chinese market, 90% were small cars. The development of electric crossovers and SUVs is, however, observed on the European market.

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

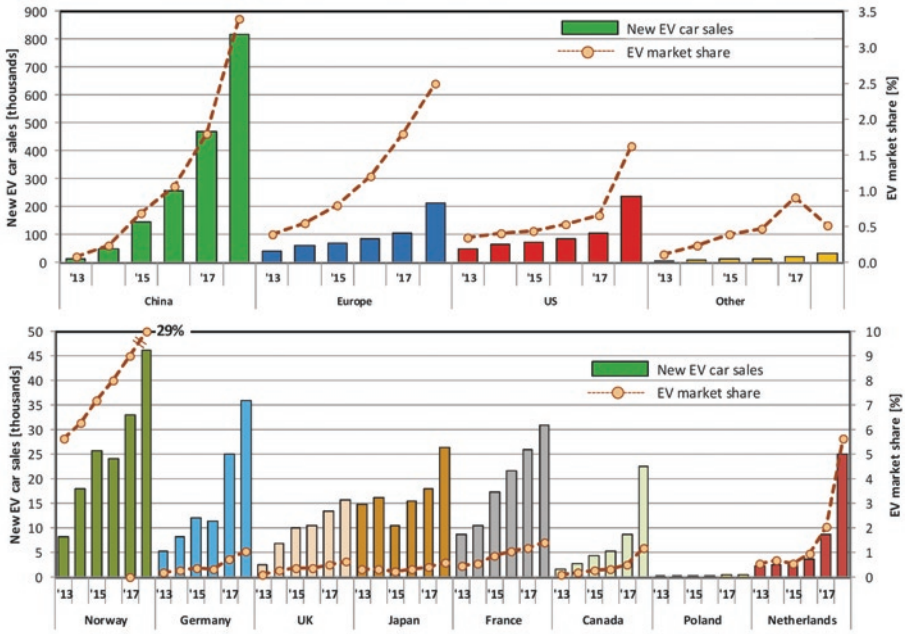


Fig. 1. Worldwide sales of electric vehicles and their share in the markets of selected countries and regions [5]

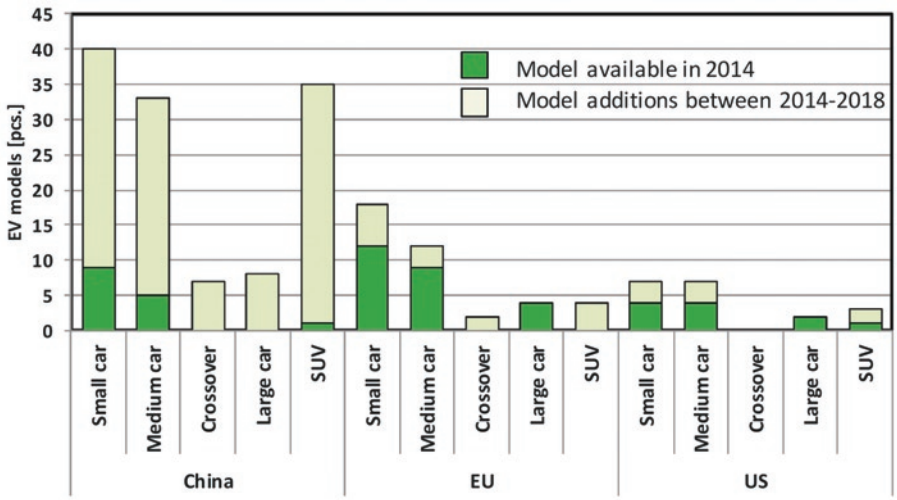


Fig. 2. Electric vehicle models divided into categories available on the markets of selected regions [5]

There are 20 BEV models officially available on the Polish market (including vans) [14]. Their range is quite diverse and is between 100 km to 540 km (in the NEDC test – New European Driving Cycle), this range is provided by batteries with capacities between 6.1 kWh and 90 kWh, respectively. The list does not include Tesla vehicles, as their official sales (on-line) only began at the end of August 2019.

The Transport & Environment [10] estimates that the share of electric vehicles in 2025 will reach around 8% of total sales and nearly 17% by 2030.

The analysis of energy consumption in conventional drive systems is based on the fuel consumption values or the carbon dioxide emissions. However, exhaust emission tests of motor vehicle drives indicate large discrepancies (between 30–40%) in fuel consumption and CO_2 emissions between type approval tests and the real driving conditions [4].

The carbon dioxide emission values are also influenced by the vehicle operating conditions as well as any equipment designed to reduce fuel consumption at a standstill. The impact of various test routes and the use of the start-stop system were the subject of research conducted during real-world traffic tests for passenger cars [8]. It was stated there that the non-repeatable nature of the vehicle operating conditions on the same test route may cause a difference in the value of carbon dioxide emissions of about 26%. Additionally, the use of the start stop system makes it possible to reduce this emission by a further 11–15%.

The impact of dynamic conditions in road tests on carbon dioxide emissions was also assessed [6]. These tests, known as RDC (Real Driving Conditions), have been required since 2018 as a part of the vehicle type approval process. The authors concluded that there is a strong relationship between carbon dioxide road emissions and dynamic driving conditions. It has been shown that doubling the value of relative positive acceleration results in a 3-fold increase in carbon dioxide road emissions.

According to Pavlovic et al. [11] CO_2 emissions in the WLTC test are approximately 10% higher than those in the NEDC test (for vehicles with SI engines with a curb weight of approximately 1500 kg), while energy consumption is increased by approximately 40% (for vehicles with SI engines).

Simulation tests of fuel consumption and CO_2 emissions using AVL Cruise conducted by Tsokolis et al. [16] indicate differences in values obtained from NEDC and WLTC tests. These differences were observed for over 63% of tested vehicles with SI engines and for 81% of vehicles with CI engines. CO_2 emissions were on average 11% higher for all vehicles in the WLTC test compared to the NEDC test. The average efficiency values were higher in the WLTC test (than in NEDC) and they amounted to: engine efficiency – 31% (compared to 25%) and vehicle efficiency – 26% (compared to 21%).

Electric vehicle propulsion analysis is currently considered in two forms. The first concerns vehicle tests, the other – simulation tests. Both test variants can be combined to determine the energy consumption of an electric vehicle.

Analyzes conducted by Wu et al. [18] indicate a higher efficiency of electric drive system in urban traffic than in the conditions of highway traffic. This is due to the greater potential for energy recovery in urban driving conditions.

The analysis of energy flow in hybrid vehicles in real traffic had been carried out for several years now [12, 13, 17]. The analysis of such drive systems is based primarily on the analysis of electrical energy use by vehicles while excluding fuel consumption. This is due to the fact that the internal combustion engine works in part as an electric power generator, which allows increasing the energy capacity of the vehicle's battery.

Current research on the electric vehicle energy consumption reduction relates to analyzes towards the optimization of the electric motors torque [19], limiting the losses generated by electric motors [15], the vehicles speed profile [1], the problem of route selection in the aspect of charging stations [7, 20], and optimization of the charging infrastructure [2, 9].

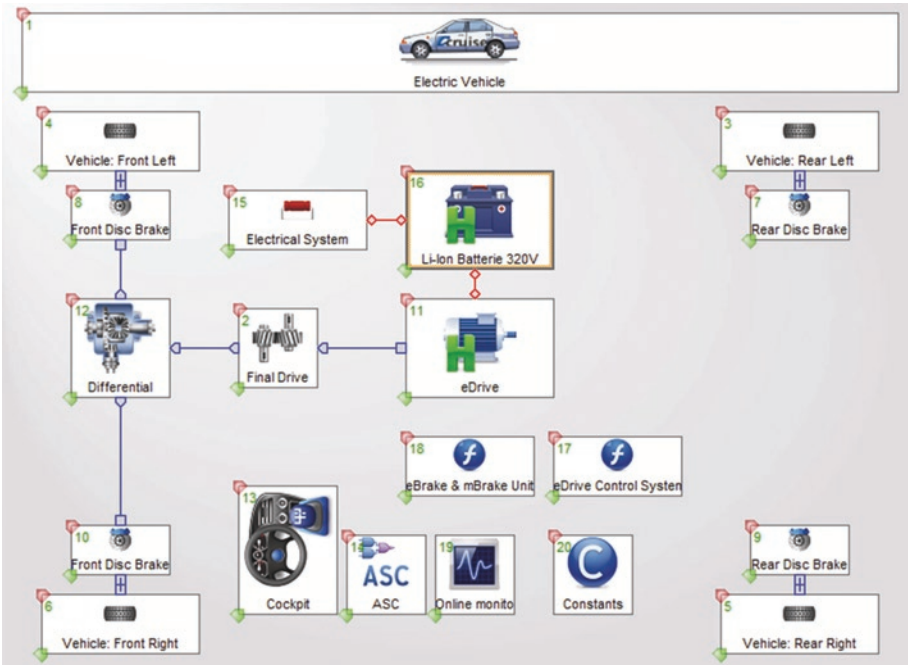


Fig. 3. Diagram of the electric vehicle drive system (AVL Cruise)

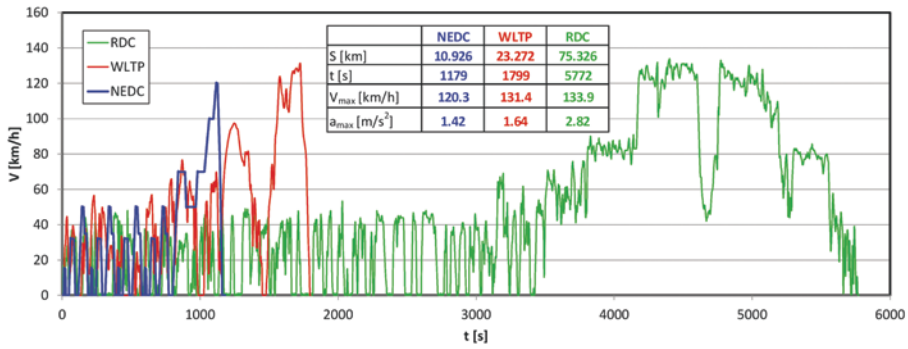


Fig. 4. Driving tests used in the research and their characteristics

The main aim of the research is to determine the differences in specific electricity consumption of motor vehicles in driving tests. It is therefore necessary to answer the question: to what extent different tests influence the estimation of specific energy consumption and whether it is possible to use such tests interchangeably. An additional factor taken into account in the tests was the variable vehicle mass.

2. Testing method

The research was carried out using AVL Cruise simulation software. It is a system enabling the simulation of a drive system (BEV, HEV or conventional), whose special features include: (1) the division of the drive system into functional elements with predefined characteristics of individual components; (2) the model structure and solution algorithm being independent; (3) generation of equations based on full system definition and (4) multi-threaded data models integrity.

An electric vehicle with variable curb weight was modeled, the structural model of which was shown in Fig. 3. The technical characteristics of the vehicle were shown in Table 1.

Energy consumption tests were carried out in relation to various driving tests (NEDC, WLTC and RDC) and for a vehicle with different curb weights (1000; 1500 and 2000 kg). The characteristics of the driving tests are shown in Fig. 4. The tests included the current NEDC driving test, the modern WLTC driving test and the RDC road test. These are characterized by a varied driving profile, different test route lengths and similar maximum driving speeds. Due to the different acceleration curves, different total energy consumption is expected.

3. Vehicle electric drive characteristics

The electric drive tests were carried out at a constant initial value of the battery charge level (SOC = 95%). Because recharging was only possible during regenerative braking, this value has not been changed. It was found that the type of driving test used significantly affects the final battery state of charge (Fig. 5). The shortest test (NEDC) causes a few percent change in the battery SOC. Increasing the length and intensity of the test results in lower final SOC values. Taking into account the length of the driving test, the final SOC value per 100 km of the test distance was determined. The final value ΔSOC/100 km in the NEDC test (taking into account the initial battery charge of 95%) was 53%, while in the WLTC test it was 42%, and in the RDC test – 45%. This means that the SOC final values were not the smallest in most aggressive RDC driving test. The maximum ΔSOC variations between the tests were 11% (NEDC and WLTC) and the values were small when comparing the ΔSOC of the WLTC and RDC tests – 3%.

In addition, an increase in vehicle curb weight causes a 5% change in SOC (NEDC test), and a 12% change for the WLTC test. During the RDC test, the vehicle weight results in the most significant changes the final SOC values. An increase in weight of 100 kg reduces battery state of charge by approximately 4% on average. This value is very important,

Table 1. Technical data of the simulated vehicle and drive system

Vehicle		
Curb weight	1000; 1500; 2000	kg
Wheelbase	2467	mm
Electric circuit		
Battery	Li-Ion, 25 kWh	
Nominal voltage	360	V
Cell capacity × number of cells	36 × 2	Ah
Electric motor		
Type	asynchronous	
Torque	240 @ 0–3000	Nm@rpm
Transmission		
Gear ratio	6.058	

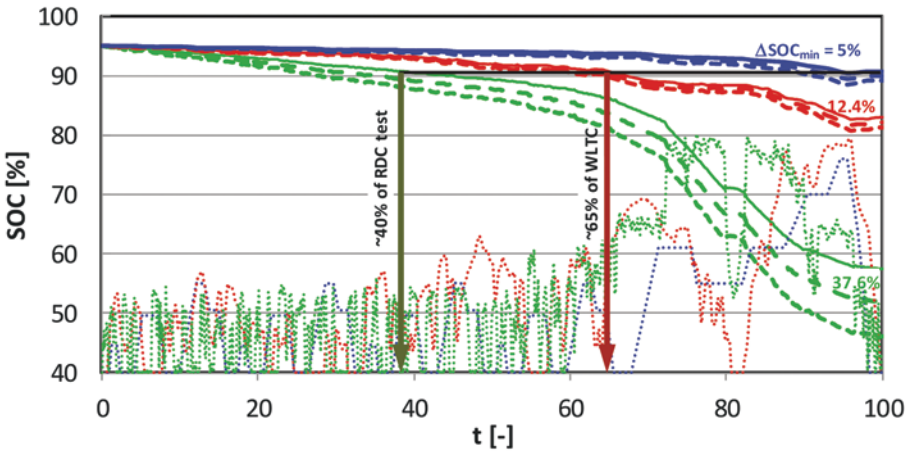


Fig. 5. Analysis of changes in the battery state of charge for different vehicle curb weights (1000 kg – solid line, 1500 kg – long dashed line, 2000 kg – short dashed line) and various driving tests)

because with a minimum vehicle weight (1000 kg) the change in the SOC value in this test was already as high as 95 – 57 = 38%. Considering that the total energy of the battery is 25 kWh and its operating capacity lies in the real SOC range of 20–80%, the change in the battery's SOC was already 63% of its entire usable energy. This means that 1/3 of the SOC changes were responsible for 2/3 of the effective energy of a Li-Ion battery.

In addition, it was found that the maximum change in SOC obtained in the NEDC test occurs after just 40% of the RDC test length and 65% of the WLTC test duration. This means that the driving test selection greatly influences the final SOC value of the battery.

To assess the range of used operating parameters of the electric motors, a fixed matrix was used in the coordinates Mo = f(n) under the following assumptions (a simplified diagram is shown in Fig. 6a):

- n = 250 rpm – resulting in 30 intervals (in the range of 0–7000 rpm),
- Mo = 20 Nm – resulting in 25 intervals (in the range of –250–250 Nm).

An example map of the electric motor operating points in the RDC test for a vehicle mass m = 2000 kg is shown in Fig. 6b.

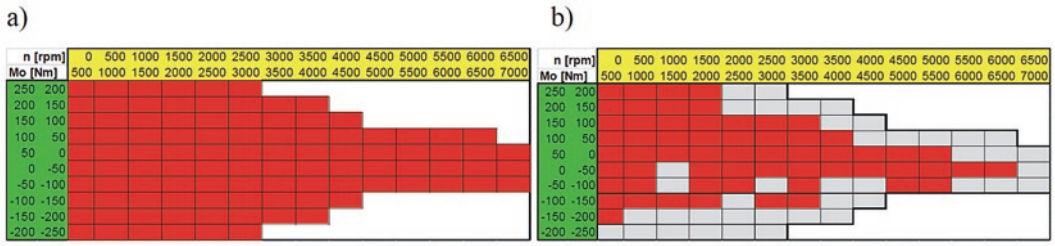


Fig. 6. Example methodology for determining the useful parameter ranges of the electric motor operation: a) full operating area, b) actual operating area of the electric motor in the RDC test

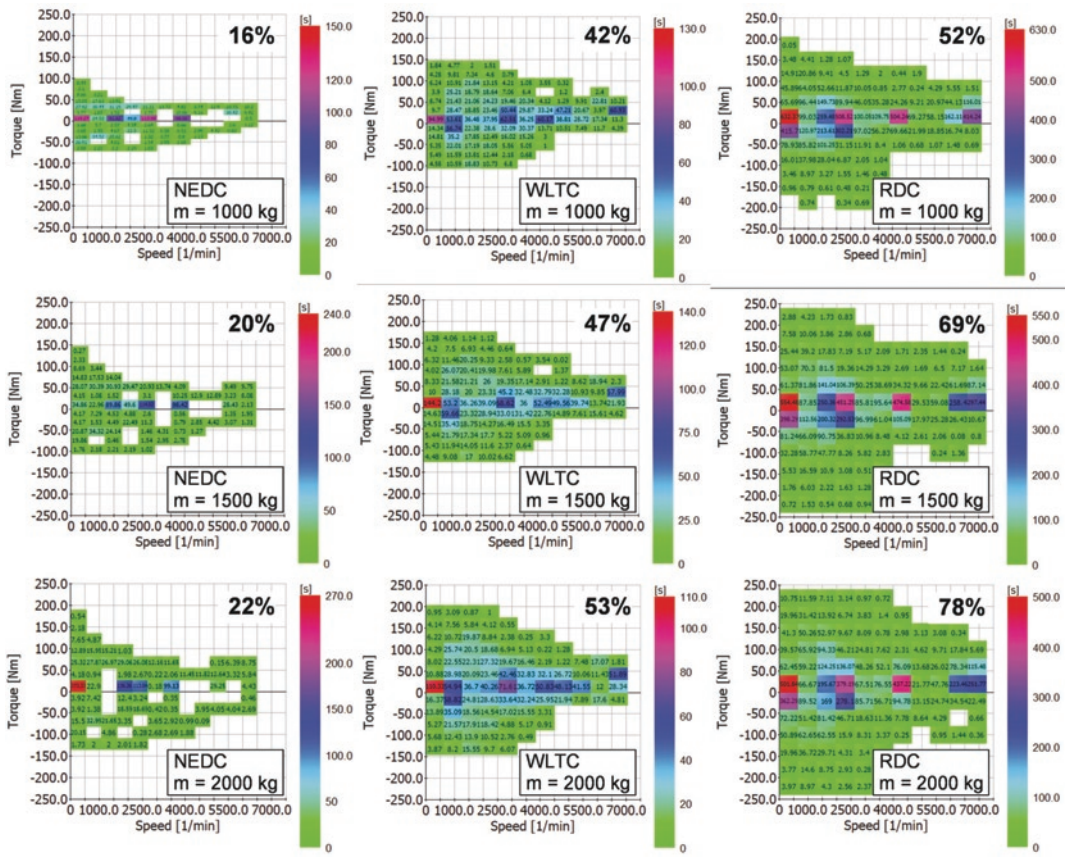


Fig. 7. Electric motor performance characteristics in various driving tests and for different vehicle curb weights

A high density of maximum torque at low speeds can be observed, as well as unused operating ranges with high engine efficiency. Regenerative braking also allows the use of a large operating range (two-quadrant operation of the electric motor), while the maximum values are limited especially in the middle range of the rotational speed of this motor.

The degree to which the full characteristic of the electric motor's operating range was used was determined for such maps. The values for the operating intervals usage were given in Fig. 7. Each operating interval contains information on the amount of time that a given interval, described by the values Mo and n, was used. At the same time, these values were in line with the legend presented for each characteristic. The main percentage values on the top right indicate the percentage of the total available range of the electric motor operating parameters that was used in any given test. An increase in the use of the electric motor's full range of operation for different drive tests can be observed. For the NEDC test, the range of motor operating parameters used was at most 22% of the total range available for the electric motor, while for the RDC test – 78% (at vehicle mass m = 2000 kg). Increasing the torque of the electric motor is most important when increasing the weight of the vehicle in a test simulating real driving conditions (RDC test).

4. Analysis of energy flow through the battery

The high-voltage battery with a nominal voltage of 360 V was simulated as a system consisting of two rows of cells with an electrical capacity of 36 Ah. This arrangement results in the total value of accumulated energy being about 25 kWh.

Energy flow simulation of was made for all test routes, with an example of battery charging and discharging process given in Fig. 8. Due to the high values of SOC changes (Fig. 5), the presented results are for the RDC test. The data shows that it is possible to recover only a dozen or so percent of the battery charge while travelling the test route.

Although the RDC test generates the largest SOC changes, it does not result in the recovery of large amounts of electric energy. Relative to NEDC and WLTC tests, the average share of battery energy from charging were only about 3–4 percentage points higher for each vehicle mass tested.

Detailed data regarding the energy flow was shown in Fig. 9. The conditions for charging and discharging the batteries as well as their total energy change were determined. The data provided shows that the increase in vehicle mass, despite the increase in the amount of energy recovered, reduces its share in the total energy balance. This result indicates the need to optimize the vehicle's mass in the context of the energy flow in the electric vehicle's drive system. Larger energy flows during the RDC test also point to the need to provide large energy capacities for BEV batteries and large initial SOC values. The total energy change values were the largest for the RDC test and were 3–4 times the energy change amount observed in the WLTC test. The differences between the energy changes in the NEDC and RDC tests were about 8 times (in favor of the NEDC test). An increase in vehicle weight (by 1/3) resulted in approximately 12% increase of the overall energy cosumed, regardless of the test used.

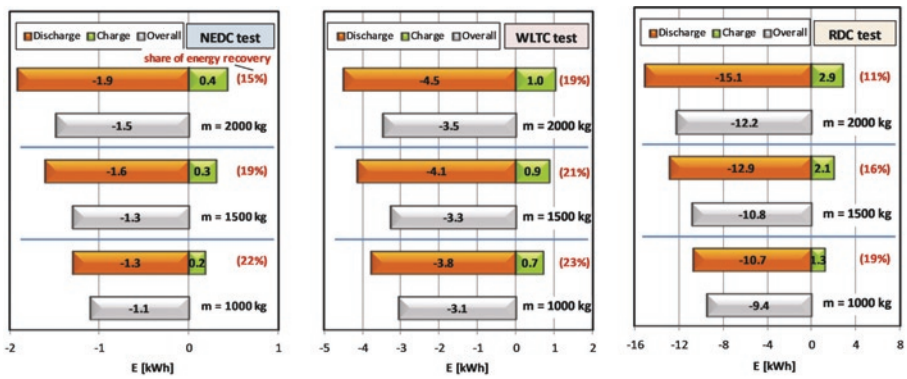


Fig. 9. Battery operating conditions and energy flow in driving tests

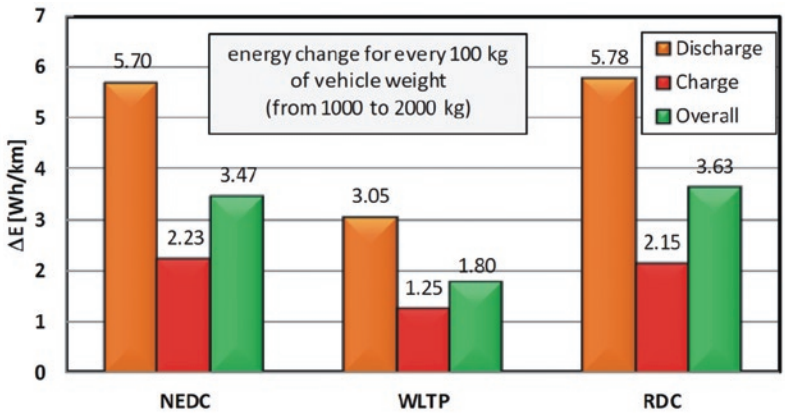


Fig. 10. Analysis of changes in energy flow through the battery, taking into account the change in vehicle mass by Δm = 100 kg

The smallest effect of mass on energy consumption was obtained in the NEDC test (about 6%), the largest in the RDC test – 17%.

The assessment of energy demand in the form of energy per 100 kg of vehicle mass and to the driving test distance was determined using the equation:

$$\Delta E = \frac{E_{i,m_2} - E_{i,m_1}}{n \cdot S_{test,j}} \tag{1}$$

where: E – battery energy value [Wh]; i – driving phases taking into account discharge, charging and total energy flow; j – driving tests carried out: NEDC, WLTC, RDC, S – driving test distance [km]; m_2 = 2000 kg, m_1 = 1000 kg; n = 10 – means the energy per 100 kg of vehicle weight.

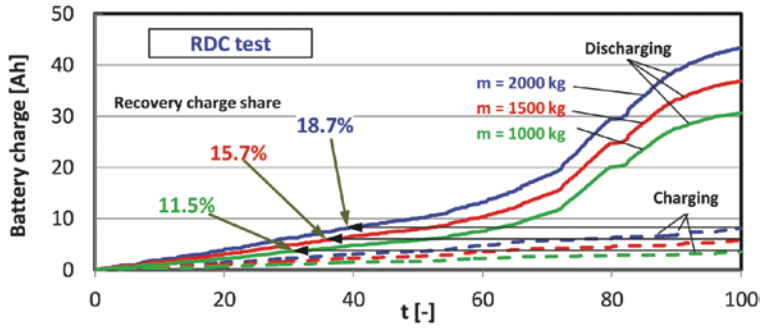


Fig. 8. Changes in the battery charge during the RDC test for different vehicle weights, and the characteristics of its charging and discharging

The results of these calculations were shown in Fig. 10. The analysis shows that similar energy expenditure per 100 kg of vehicle mass and per unit test length occur for NEDC and RDC tests. Although the NEDC test is less dynamic compared to WLTC, the specific energy expenditure is much higher (by about 50%) according to equation (2). It should be noted that such parallels are similar in NEDC and RDC tests; these similarities occur both when discharging and charging the batteries.

5. Analysis of the vehicle's specific energy consumption

Vehicle energy consumption was determined in the form of specific energy for each of the driving tests and taking into account the weight of the vehicle in the range from 1000 kg to 2000 kg. Specific energy consumption [kWh/100 km] was defined as the amount of energy consumed to perform a test drive length of 100 km:

$$E_j = \frac{\int_{t=0}^{t} E dt}{S} \cdot 100 \tag{2}$$

where: t_0 – is the start of the test, t_{max} – test duration, S – test drive distance E – the amount of energy used in the test at a given time.

This energy was determined in two ways: in the whole test (E_{end}) and as a function of time during the test – E_j (Fig. 11). The initial phase of each test generates large E_j values (due to the short duration of the test), this value stabilizes in further test phases. The amount of time it takes for the final energy consumption value in the test to stabilize depends on the dynamics of the driving test. The specificity of driving tests means that even during the static NEDC test, the final energy consumption value was obtained after 98% of its duration. It was assumed that the energy consumption value will have reached its final value when it falls within the range of 5% from the value observed by the end of the test:

$$t_{\%} = \pm 5\% \text{ of } E_{end} \tag{3}$$

where: $t_{\%}$ – relative stabilization time (within ±5% of the final determined total energy consumption value – E_{end}).

During the analysis of the more dynamic WLTC and RDC tests, $t_{\%}$ values were reached at 98% and 90% of the total test duration, respectively (Fig. 11a). This means that for the most dynamic RDC test, the time to determine the final value of total energy consumption was the shortest (Fig. 11c). It should be noted that the modification of the test phases (urban, extra-urban, highway driving) may contribute to earlier determination of the final value (assuming that the order of the driving phases does not change).

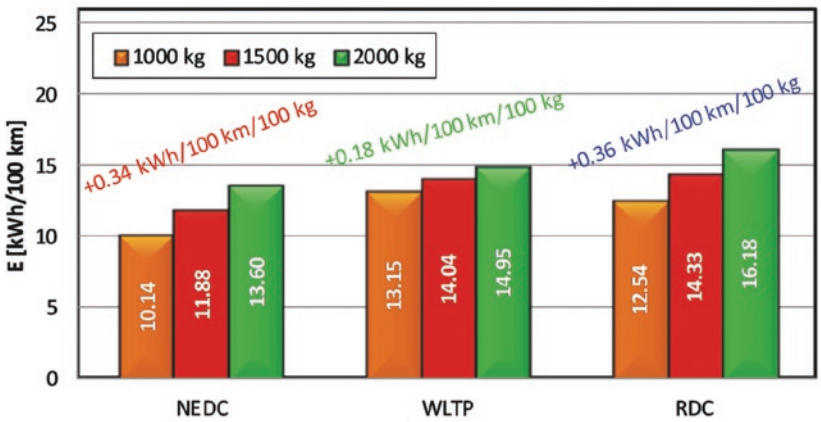


Fig. 12. Vehicle energy consumption values in driving tests, taking into account the changes in vehicle mass

The shortest energy consumption value stabilization time was obtained in the RDC test. In this (most dynamic) test, the final energy consumption value was also the highest. The lowest values of energy consumption occurred in the NEDC test. A proportional effect of the vehicle mass on the final energy consumption values of each test should be stated.

The total energy consumption analysis results were shown in Fig. 12. The highest values of this consumption were obtained in the RDC test. They were measured to be about 20% higher than the energy consumption value in the NEDC test. The specific increases in this energy consumption also related to the increase in vehicle weight by 100 kg were also indicated there. It was found that the increases in energy consumption were the same for NEDC and RDC tests and they were about 0.34 kWh/100 km for each additional 100 kg of vehicle weight. The energy consumption change depending on the weight of vehicles recorded for the WLTC test increased by only half the value for the other two tests.

Such estimation of the energy consumption increase value allows to determine this consumption without the need to perform simulations or real tests of the vehicle with its own weight changed.

6. Conclusions

Based on the simulation tests (using AVL Cruise software) the operating conditions of EV motors and the energy consumption of the vehicle with different curb weight values were determined in various driving tests. Based on them, the following conclusions were formulated:

1. Regarding the level of battery charge:
   - The dynamic character of the RDC test resulted in the largest changes in battery state of charge (ΔSOC). These changes are three times higher than in the WLTC test and 7 times higher than in the NEDC test.

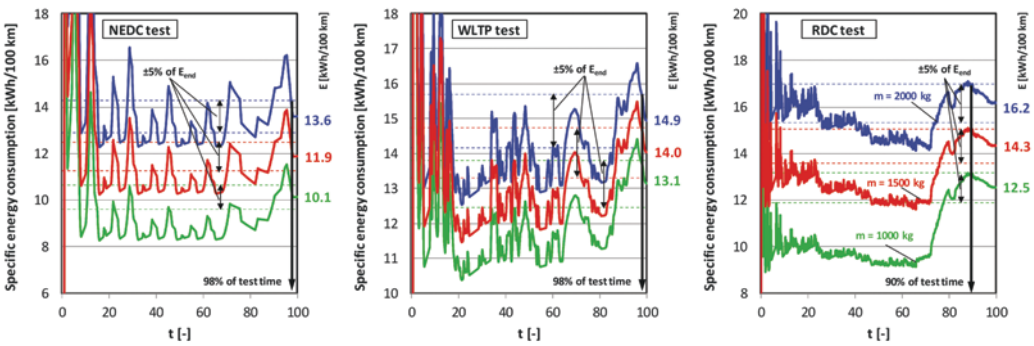


Fig. 11. Conditions for determining the final energy consumption value of the driving tests

- Taking into account the test length means that the ratio of changes in ΔSOC per 100 km of the tests RDC; WLTC; and NEDC were: 1.06; 1; 1.25 respectively. This means a change in ΔSOC/100 km in percentage values of: 53%: 42%: 45% respectively. The smallest differences were recorded in the WLTC and RDC tests – up to only 6%.

2. Regarding the energy flow through the battery:
   - The values for the total change in battery charge and discharge energy were highest in the RDC test; they were 3–4 times greater than the energy changes measured in the WLTC test. The energy changes in the NEDC and RDC tests differed by about 8 times (in favor of the NEDC test).
   - An increase in vehicle weight (by 1/3) results in an approximately 12% increase in the overall energy consumption, regardless of the test used. The smallest effect of mass on the total energy flow through the battery was obtained in the NEDC test (about 6%), the largest in the RDC test – 17%.
3. Regarding the energy consumption in the test:
   - The highest energy consumption values were obtained during the RDC test. They were about 20% greater than the energy consumption value in the NEDC test. During the WLTC test, the energy consumption values were 30 to 10% greater than the energy consumption in the NEDC test (corresponding to increasing the weight of the vehicle).
   - The increase in energy consumption in the NEDC and RDC tests was the same and amounts to approximately 0.34 kWh/100 km for each additional 100 kg of vehicle weight. The increases in energy consumption depending from increased weight for the WLTC test were recorded to be only half that value.

The obtained tests and analyzes results of electric vehicles energy consumption indicate the need for further simulation works and tests in real driving conditions. The energy consumption of such vehicles was strongly dependent on the test type; values obtained in driving tests were at the level of 10.1–13.5 kWh/100 km (NEDC test); 13–15 kWh/100 km (WLTC test) and 12.5–16.2 kWh/100 km in the RDC test. The differences in controled driving tests (NEDC and WLTC) reached up to 25% (larger in the WLTC test for vehicles with a lower curb weight). Tests conducted in real conditions show similar values of energy consumption (for light vehicles – about 1000 kg) and an increase of this consumption by another 10% when vehicles with a mass of about 2000 kg were tested.

Acknowledgements

The research was conducted using the AVL Cruise software thanks to the AVL University Partnership Program.

References

1. Basso R, Kulcsár B, Egardt B, Lindroth P, Sanchez-Diaz I. Energy consumption estimation integrated into the Electric Vehicle Routing Problem. Transportation Research Part D: Transport and Environment 2019; 69: 141–167, https://doi.org/10.1016/j.trd.2019.01.006.
2. Davidov S, Pantoš M. Planning of electric vehicle infrastructure based on charging reliability and quality of service. Energy 2017; 118: 1156–1167, https://doi.org/10.1016/j.energy.2016.10.142.
3. European Commission. Proposal for a regulation of the European Parliament and of the Council setting emission performance standards for new passenger cars and for new light-commercial vehicles as part of the Union's integrated approach to reduce CO2 emissions from LDVs. Brussels, 8.11.2017, SWD(2017) 650 final. ec.europa.eu/clima/sites/clima/files/transport/vehicles/docs/swd_2017_650_p1_en.pdf.
4. Fontaras G, Zacharof N-G, Ciuffo B. Fuel consumption and CO2 emissions from passenger cars in Europe – Laboratory versus real-world emissions. Progress in Energy and Combustion Science 2017; 60: 97–131, https://doi.org/10.1016/j.pecs.2016.12.004.
5. IEA, Global EV Outlook 2019. IEA, Paris, www.iea.org/publications/reports/globalevoutlook2019.
6. Kurtyka K, Pielecha J. The evaluation of exhaust emission in RDE tests including dynamic driving conditions. Transportation Research Procedia 2019; 40: 338–345, https://doi.org/10.1016/j.trpro.2019.07.050.
7. Langbroek J H M, Cebecauer M, Malmsten J, Franklin J P, Susilo Y O, Georén P. Electric vehicle rental and electric vehicle adoption. Research in Transportation Economics 2019; 73: 72–82, https://doi.org/10.1016/j.retrec.2019.02.002.
8. Merkisz J, Pielecha J, Radzimirski S. New trends in emission control in the European Union. Springer Tracts on Transportation and Traffic 2014; 4: 170, https://doi.org/10.1007/978-3-319-02705-0.
9. Micari S, Polimeni A, Napoli G, Andaloro L, Antonucci V. Electric vehicle charging infrastructure planning in a road network. Renewable and Sustainable Energy Reviews 2017; 80: 98–108, https://doi.org/10.1016/j.rser.2017.05.022.
10. Muzi N. New car CO2 standards: Is the job of securing electric cars in Europe done? Transport & Environment 2019. www.transportenvironment.org.
11. Pavlovic J, Marotta A, Ciuffo B. CO2 emissions and energy demands of vehicles tested under the NEDC and the new WLTP type approval test procedures. Applied Energy 2016; 177: 661–670, https://doi.org/10.1016/j.apenergy.2016.05.110.
12. Pielecha I, Cieslik W, Szalek A. Operation of electric hybrid drive systems in varied driving conditions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (1): 16–23, https://doi.org/10.17531/ein.2018.1.3.
13. Pielecha I, Cieslik W, Szalek A. Operation of hybrid propulsion systems in conditions of increased supply voltage. International Journal of Precision Engineering and Manufacturing 2017; 18: 1633–1639, https://doi.org/10.1007/s12541-017-0192-3.
14. PSPA, 2019. Licznik elektromobilności. Polskie Stowarzyszenie Paliw Alternatywnych. pspa.com.pl
15. Sun B, Zhang T, Ge W, Tan C, Gao S. Driving energy management of front-and-rear-motor-drive electric vehicle based on hybrid radial basis function. Archives of Transport 2019; 49 (1): 47–58, https://doi.org/10.5604/01.3001.0013.2775.
16. Tsokolis D, Tsiakmakis S, Dimaratos A, Fontaras G, Pistikopoulos P, Ciuffo B, Samaras Z. Fuel consumption and CO2 emissions of passenger cars over the New Worldwide Harmonized Test Protocol. Applied Energy 2016; 179: 1152–1165, https://doi.org/10.1016/j.apenergy.2016.07.091.
17. Wei Z, Xu Z, Halim D. Study of HEV power management control strategy based on driving pattern recognition. Energy Procedia 2016; 88: 847–853. https://doi.org/10.1016/j.egypro.2016.06.062.
18. Wu W, Freese D, Cabrera A, Kitch W A. Electric vehicles' energy consumption measurement and estimation. Transportation Research Part D: Transport and Environment 2015; 34: 52–67, https://doi.org/10.1016/j.trd.2014.10.007.
19. Xie L, Luo Y, Zhang D, Chen R, Li K. Intelligent energy-saving control strategy for electric vehicle based on preceding vehicle movement. Mechanical Systems and Signal Processing 2019; 130: 484–501. https://doi.org/10.1016/j.ymssp.2019.05.027.

20. Zhang S, Gajpal Y, Appadoo S S, Abdulkader M M S. Electric vehicle routing problem with recharging stations for minimizing energy consumption. International Journal of Production Economics 2018; 203: 404–413, https://doi.org/10.1016/j.ijpe.2018.07.016.

Ireneusz PIELECHA
Jacek PIELECHA
Institute of Combustion Engines and Transport
Poznan University of Technology
ul. Piotrowo 3, 60-965 Poznań, Poland

E-mail: Ireneusz.Pielecha@put.poznan.pl,
Jacek.Pielecha@put.poznan.pl

Article citation info:
KŁOSOWSKI G, RYMARCZYK T, KANIA K, ŚWIĆ A, CIEPLAK T. Maintenance of industrial reactors supported by deep learning driven ultrasound tomography. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 138–147, http://dx.doi.org/10.17531/ein.2020.1.16.

Grzegorz KŁOSOWSKI
Tomasz RYMARCZYK
Konrad KANIA
Antoni ŚWIĆ
Tomasz CIEPLAK

MAINTENANCE OF INDUSTRIAL REACTORS SUPPORTED BY DEEP LEARNING DRIVEN ULTRASOUND TOMOGRAPHY

EKSPLOATACJA REAKTORÓW PRZEMYSŁOWYCH ZE WSPOMAGANIEM TOMOGRAFII ULTRADŹWIĘKOWEJ I ALGORYTMÓW GŁĘBOKIEGO UCZENIA

Monitoring of industrial processes is an important element ensuring the proper maintenance of equipment and high level of processes reliability. The presented research concerns the application of the deep learning method in the field of ultrasound tomography (UST). A novel algorithm that uses simultaneously multiple classification convolutional neural networks (CNNs) to generate monochrome 2D images was developed. In order to meet a compromise between the number of the networks and the number of all possible outcomes of a single network, it was proposed to divide the output image into 4-pixel clusters. Therefore, the number of required CNNs has been reduced fourfold and there are 16 distinct outcomes from single network. The new algorithm was first verified using simulation data and then tested on real data. The accuracy of image reconstruction exceeded 95%. The results obtained by using the new CNN clustered algorithm were compared with five popular machine learning algorithms: shallow Artificial Neural Network, Linear Support Vector Machine, Classification Tree, Medium k-Nearest Neighbor classification and Naive Bayes. Based on this comparison, it was found that the newly developed method of multiple convolutional neural networks (MCNN) generates the highest quality images.

Keywords: *deep learning; inverse problem; ultrasound tomography; image reconstruction; process tomography.*

Monitorowanie procesów przemysłowych jest ważnym elementem zapewniającym właściwą eksploatację urządzeń i wysoki poziom niezawodności procesów. Prezentowane badania dotyczą zastosowania metod głębokiego uczenia w obszarze eksploatacji zbiornikowych reaktorów przemysłowych. W procesach przemysłowych opartych na reakcjach chemicznych zachodzących wewnątrz procesowej tomografii ultradźwiękowej (UST). Opracowano nowatorski algorytm wykorzystujący jednocześnie wiele klasyfikacyjnych splotowych sieci neuronowych (CNN) do generowania monochromatycznych obrazów 2D. Aby osiągnąć kompromis między liczbą sieci a liczbą wszystkich możliwych wyników pojedynczej sieci, zaproponowano podział obrazu wyjściowego na klastry 4-pikselowe. W związku z tym liczba wymaganych CNN została czterokrotnie zmniejszona i istnieje 16 różnych wyników z jednej sieci. Nowy algorytm został najpierw zweryfikowany przy użyciu danych symulacyjnych, a następnie przetestowany na danych rzeczywistych. Dokładność rekonstrukcji obrazu przekroczyła 95%. Wyniki uzyskane przy użyciu nowego algorytmu klastrowego CNN zostały porównane z pięcioma popularnymi algorytmami uczenia maszynowego: płytką sztuczną siecią neuronową, maszyną liniowego wektora wsparcia, drzewem klasyfikacji, klasyfikacją średniego k-najbliższego sąsiada i naiwnym Bayesem. Na podstawie tego porównania stwierdzono, że nowo opracowana metoda wielu splotowych sieci neuronowych (MCNN) generuje obrazy o najwyższej jakości.

Słowa kluczowe: *uczenie głębokie; problem odwrotny; tomografia ultradźwiękowa; rekonstrukcja obrazu; tomografia procesowa.*

1. Introduction

1.1. Tank chemical reactors

The basic device for the implementation of the batch processes are tank reactors. Batch processes are widely used in many branches of economy e.g. food, chemistry, pharmacy, semiconductors, biogas plants and so on. Due to the time-varying, non-linear and uneven nature of this process, it is very difficult to determine the exact mathematical model of these processes, which necessitates their monitoring. For this reason, to ensure a high level of reliability and trouble-free maintenance of tank reactors it is necessary to effectively monitor the processes taking place inside them.

A tank chemical reactor is, in the simplest sense, a vessel adapted to carry out a specific chemical reaction in it. On an industrial scale, the construction of a reactor and the parameters of its process should ensure optimal economic results. Chemical reactors and the processes taking place in them are usually an essential element of a technological process aimed at producing a specific chemical product. Any other processes in such a sequence should be assigned a rather auxiliary role, consisting either in preparatory activities or in separating the products of the reaction and separating from them the component, the obtaining of which is the aim of production operations. Produc-

tion and economic results usually depend on the correct operation of a chemical reactor. The chemical process occurring in the reactor overlaps, to varying degrees and in different proportions, the processes of mass, momentum and heat transfer. This usually gives a very complicated picture of the total process. The description of these total complex processes can often be simplified if one of the mentioned processes plays a dominant role in it and the others can be omitted. The most often such a dominant elementary process is the transport of mass or energy. Effective monitoring of these types of processes requires non-invasive methods that include tomography.

In the reactor classification prepared according to the type of reagent cluster state, we distinguish two basic reactor groups - homogeneous and heterogeneous. In heterogeneous reactors, gas and solids, gas and liquid, solids and liquids can react with each other. The presented studies concern the monitoring of industrial two-phase processes of solids and liquids (crystallization) and liquids and gases (detection of gas bubbles in liquids or suspensions).

A specific type of chemical reactors are bioreactors. Bioreactors comply with the same rules as chemical reactors. The difference, however, lies in the complexity of the system which is a living organism. Biological reactions are more sensitive and less stable, and therefore require more attention to process control and analysis of more factors than chemical reactions. In order to ensure the proper operation of bioreactors, it is necessary to use an effective method of monitoring industrial processes. In this context, the issues of monitoring biological processes deserve attention [1,24,34].

To rationally control the maintenance processes of technical facilities, including reactors, it is necessary to know about the state of these objects. To this end, various monitoring methods are used. Monitoring allows you to solve important maintenance problems, which include maximizing reliability and process diagnostics [40]. Tomography is a cheap and non-invasive method of monitoring, but in industry, this method does not always allow obtaining high resolution images. For this reason, research is needed that will lead to the method of industrial tomography, which will guarantee high-resolution cross-sectional images of the examined objects. This article refers to a new method of ultrasonic tomography using deep learning algorithms [19].

1.2. Tomographic methods and algorithms

Tomography is used in the areas where obtaining good image quality involves the use of a non-invasive monitoring method [9]. There are many types of tomographic methods, including: computed tomography (CT) [3,29], radio tomographic imaging (RTI) [5], electrical impedance tomography (EIT) [30,31], electrical capacitance tomography (ECT) [4,16,17,22,42] or ultrasound tomography (UST) [45].

UST enables non-invasive visualization of the interior of the tested object based on measurements of the time of sound propagation between different points on the perimeter of the tank. The presence of inclusions in the object changes the time of fly of sound waves. On this basis, the cross-sectional image of the tested tank can be reproduced. The main problem of UST is the difficulty of building a physical model that would be able to take into account the full complexity of acoustic phenomena occurring in a relatively small space [8]. This causes problems in implementing classic tomographic methods. This work is our first attempt to verify the hypothesis that a properly constructed neural network is able to solve the inverse problem in ultrasound tomography. Positive results obtained in this work give hope that some problems related to modeling can be bypassed by means of convolutional neural networks (CNN) [38].

Considering the number and frequency of publishing new scientific studies on innovative industrial solutions - the UST method is relatively less widespread. In the field of industrial applications, most innovations are created in the areas of electrical impedance tomography [13,32], electrical capacitance tomography (ECT) [27,28,35] and magnetic resonance imaging [23,15]. Progress is made both in the field of hardware (sensors, computer distributed systems) and in the field of reconstruction algorithms for tomographic images. The classic methods of solving the inverse problem in the tomography include Gauss Newton's (GN) method. Compared to machine learning methods, the GN method is more universal, which means that it can be used in commercial tomographs on a wider range of applications. When it comes to problems for which machine learning algorithms can be trained, the GN method is rather less precise and increasing precision of GN method requires the usage of iterative methods, which significantly slows the algorithm.

Among the statistical methods of machine learning, the following methods have been successfully applied in the reconstruction of tomographic images, especially in electrical tomography: elastic net, lasso (least absolute shrinkage and selection operator), least-angle regression (LARS) [6], k-nearest neighbors (KNN), naive Bayes, multivariate adaptive regression splines (MARS), classification tree, support vector machine (SVM), gradient boosting machine (GBM) [21], principal component or partial least square regression [22]. There is a general trend that the importance of predictive algorithms is growing in industrial applications [41,12,14,20,10].

Due to the specific conditions of the functioning of industrial systems, the use of electrical tomography is not always possible. In situations where the tested object cannot be completely isolated from the influence of other sources of electric current or when the tested environment is dielectric, ultrasound tomography (UST) can be used. There is evidence in the literature of the effective use of UST in medicine which may indicate the existence of unused potential of ultrasound process tomography [36].

Hao et al. localized the fetal abdominal standard plane from ultrasound recordings using CNN [7]. In that case, the accuracy of the system reached 90%. Zhang et al. proposed a diagnosis system based on the two-layer CNN architecture for the classification of breast tumors [43]. The accuracy of the system they developed was 93.4%. Ma et al. classified thyroid nodules based on ultrasound, using two CNNs simultaneously. The average classification efficiency was 83.02% [21]. In other medical studies, Arevalo et al. classified changes in breast cancer from hand-segmented mammography films using CNN. They managed to achieve an accuracy of 82% [2].

The above and other examples of successful combination of UST and CNN in medical applications have given rise to targeted research on effective technics for reconstructing tomography images based on the measurements of the time-of-flight of the ultrasonic waves in an industrial tank. The non-destructive UST method supported by deep learning algorithms is suitable for monitoring industrial processes occurring not only inside closed reactors but also flow processes inside complicated pipe systems.

1.3. Convolutional neural networks

CNN has in recent years become the leading topic of technical progress in many areas [37].Due to the high ability to recognize specific features visible on CNN images, they are often used in various classification problems [26,6]. Other areas of CNN application include image classification [25], video analysis, natural language processing, city monitoring, industrial cameras and all kinds of devices providing images made with ordinary photographic techniques [44]. Creating images obtained by resolving the inverse problem, such as tomographic images, is a relatively new area of application for CNN and is therefore difficult [18,11].

CNNs are good in mapping complex nonlinear functions. For this reason, the number of attempts to use deep learning algorithm for image reconstruction in EIT/ERT (electrical impedance tomography/ electrical resistance tomography) is increasing [39].

Attempts are also made to create hybrid methods, an example of which is a real-time reconstruction algorithm that produces high-quality sharp EIT absolute images by combining the D-bar algorithm with subsequent CNN processing [19].

1.4. Research objective and novelties

The main goal of the research presented in this article was to develop a new algorithm capable of reconstructing monochrome (binary) 2D images of ultrasound tomography, regardless of the size, shape, location or number of inclusions hidden in the examined object.

High accuracy of imaging is achieved thanks to splitting the outputs for multiple CNNs. The higher speed of the algorithm is reached thanks to the use of 4-pixel clusters approach. Instead of training CNN networks for all 1024 pixels, there are only 256 networks to train, one for each cluster. As a result, we get a 4-fold increase in the speed of the algorithm, preserving reasonably small number of possible outcomes of the CNNs.

Fig. 1a shows the scheme of the algorithm based on the structure of an ordinary CNN with 496 inputs and 1024 binary outputs (0, 1). In Fig. 1b, for comparison, proposed multiple CNN scheme is shown. Each of the 256 CNNs generates a numerical classifier {1..16} at the output, which is the cluster identifier. Then, each cluster is converted to a 4-pixel monochrome pattern with dimensions of 2×2 (Table 1).

This article consists of 4 sections. Section 1 presents the state of art regarding tomographic methods and algorithms used in the reconstruction of tomographic images. Specific examples of UST applications combined with deep learning are presented. Section 2 contains a detailed description of the test stand, the data used, the multiple CNN algorithm, as well as information on the learning process. Section 3 presents examples of reconstructions obtained by using the multiple CNN method. In addition, the quality of this method was compared with 5 other classical machine learning algorithms.

2. Hardware, Algorithms and Methods

2.1. The hardware

The research used a tomographic system developed by the authors of this publication. Fig. 2a shows the mesh of lines indicating measurements of the speed of the sound wave. The densities of the lines at the periphery of the tank cross section are determined by 32 transducers. Fig. 2b demonstrates the test stand. Based on measurements from this stand, effective algorithm for generating simulation data was developed.

The tomograph is built on the STM32F103VCT6 processor (Fig. 3a) which is responsible for the managing of a measurement sequence and setting transducers in the transmitting or receiving mode. Measurement data acquisition system cooperates with dedicated transducers, as shown in Fig. 3b. Measurements data can be transmitted to the PC in real time (Fig. 3c).

The transducers performs measurements using one piezoelectric unit using the absorption method. The transducer can work both as a transmitter and an ultrasonic wave receiver. The transducer has an integrated signal processing system and a microcontroller with a built-in A/C converter. By using a programmable digital potentiometer, the gain of the received signal can be regulated. The PCB transducer also provides the option of filtering out the signal using an active filter. A small diode is used to signal the operating status of the device.

Control and reading of measurements made by the transducer is carried out via the CAN 2.0A bus. Due to the special design, the

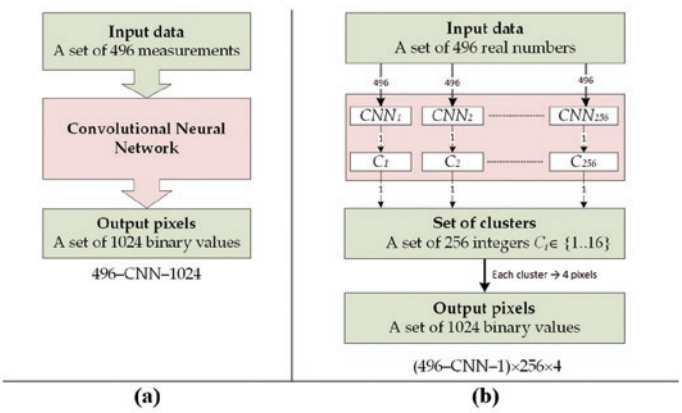


Fig. 1. Comparison of algorithms for methods CNN (a) and MCNN (b)

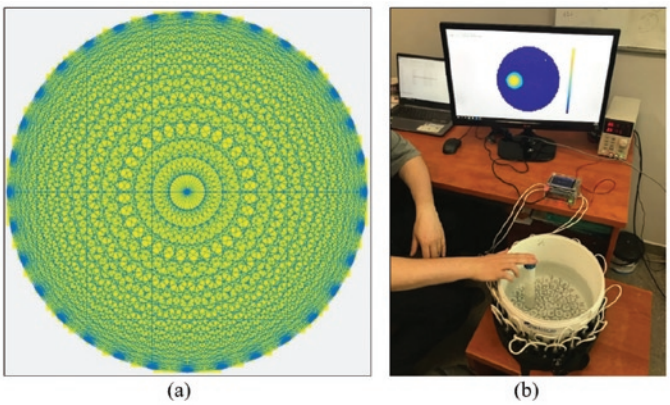


Fig. 2. (a) - A mesh of measurements carried out by 32 transducers arranged around the tank; (b) - View of the test stand [33]

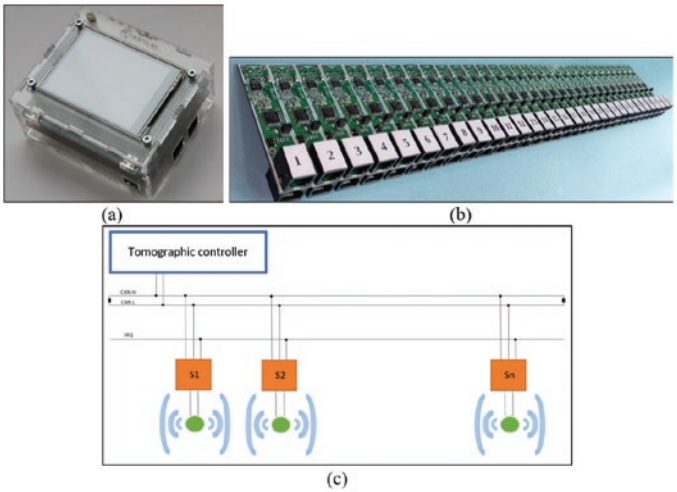


Fig. 3. (a) - Controller running on the STM32F103VCT6 processor; (b) - View of composite 32 PCBs of ultrasound transducers; (c) - Block diagram of the system

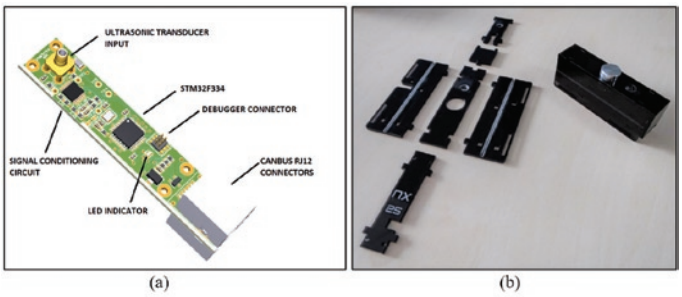


Fig. 4. Design of the active measuring transducer: (a) motherboard; (b) assembly

transducers can be installed in close proximity to each other. RJ-12 cables were used to make the communication buses and to provide power supply. Each transducer was divided into two parts - digital and analog. The task of the digital part is to send the measurement results to the tomographic controller. The analog part of the transducer was adapted to work with a 40 kHz piezoelectric unit (Fig. 4a and 4b).

2.2. Algorithms and Methods

In order to solve the inverse problem, classification convolutional neural networks (CNNs) were used. During the research, it turned out that a single convolution network powered by 496 measurements and generating 1024 binary elements learns very long and the obtained results are not good enough. The novel element of the presented solution is to use CNN to classify 4-pixel clusters.

During a single measurement series, each of the transducers acts as an emitter, while the other transducers receive ultrasound signals at that time. When the number of transducers is 32, the number of measurements can be calculated as ($32^2 - 32$). Since the time the sound wave needs to travel from A to B is the same as from B to A, the number of measurements M can be reduced by half. Hence, we get the relationship (1).

$$M = \frac{n^2 - n}{2} \tag{1}$$

where n is the number of transducers.

So the measurement vector establishing the MCNN input set that consists of 496 measurements. Each measurement means the so-called sound wave flight time between a specific pair of transducers. To ensure correct measurements, a reference measurement in an inclusion-free environment should be performed before actual measurements can begin. This is to determine the background values. Inclusions distort measured times, so that images can appear on the basis of contrast to the background. Based on different measurements of times, the locations and size of inclusions are determined.

An industrial tank filled with tap water played the role of the examined object. Various objects were immersed in water, followed by ultrasonic measurements. Thanks to the knowledge of the location and dimensions of the immersed objects, as well as the knowledge of the number of all inclusions, it was possible to develop a simulation algorithm that generated 20,000 learning cases. Fig. 5 shows a simulation example of generating one of 20,000 measurement cases.

The left side of the drawing illustrates the exemplary cross-section of the tank interior with visible inclusions. The right side is a graph of 496 measurements (horizontal axis), along with the corresponding transition times of ultrasound waves between pairs of transducers. In the simulation algorithm, Gaussian noise has been implemented with a distinct level for each measurement in the frame. The level of the noise was determined by the standard deviation set to 5% of the value of each measurement. The algorithm based on convolutional neural networks was developed using the Deep Learning Toolbox of Matlab.

As mentioned before, each learning case consisted of 496 input measurements and one monochrome output image with a resolution of 32×32. The number of measurements is the result

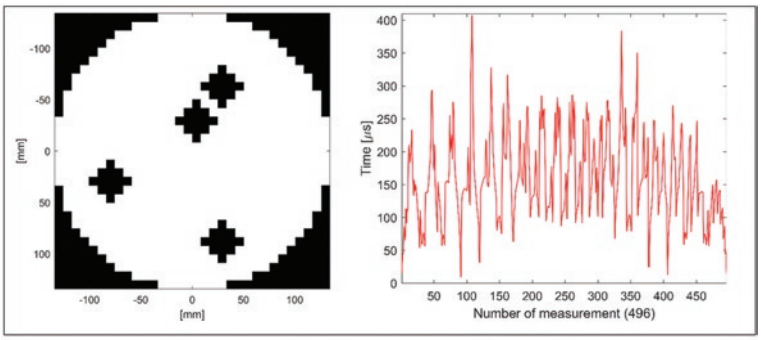


Fig. 5. A measuring case generated with the simulation method with a graph showing the times of sound waves moving between transducers

of using 32 transducers. The full matrix of measurements counts 992 (32×31) measurements, of which half of the measurements concern the same transducers. Because the sound wave moves at the same speed regardless of the direction ($v_{1-2} = v_{2-1}$), the measurement matrix tend to be symmetrical one. Therefore by averaging symmetrical measurements we obtain 496 inputs for one measurement cycle. The 32×32 monochrome image can be represented by a 1024-point vector with binary values 0 or 1. The method of converting pixels into clusters is shown in Fig. 6.

The Fig. 6 distinguishes three sample clusters with identifiers: 1, 51 and 137, as well as the method of changing pixels into the array of clusters. For example, cluster [1] contains pixels [1, 2; 33, 34]. After transforming pixels → clusters, a matrix of 1024 pixels with dimensions of 32×32 is reduced to a 16×16 matrix composed of 256 clusters.

Formally, a distinct block of four pixels from a full image can be expressed as (2):

$$B_n = \begin{bmatrix} x_n & x_{n+32} \\ x_{n+1} & x_{n+33} \end{bmatrix} \tag{2}$$

assuming that n is an odd number. This block is transformed into a cluster with a number $k(B_n)$ (3):

$$k(B_n) = \frac{1}{2}\left((n-1)\%32 + 16\frac{n-1}{32}\right) + 1 \tag{3}$$

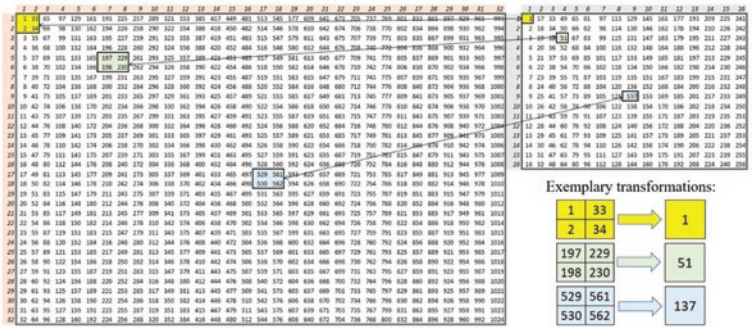


Fig. 6. The way to convert 1024 pixels into 256 clusters

Table 1. Pattern indexes p_i = {1,2,...,16}

p_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
pattern																

where $a\%b$ means the remainder of division a by b for $a=(n-1)$, $b=32$ and $\lfloor\cdot\rfloor$ is the floor function.

Since the clusters have dimensions of 2×2 that means that there are $2^4=16$ combinations of binary patterns for each cluster. The possible binary patterns of clusters are presented in Table 1.

Transforming pixels into clusters reduces number of output variables to 256. We were decided to train 256 separate CNNs, each of which generated a value from 1 to 16, corresponding to the given cluster pattern. So in this case we are talking about multiple convolutional neural networks (MCNNs), where each of the distinct CNN solves the classification problem. Table 2 shows the structure of the convolutional neural network. Each of 256 CNNs has the same structure including 15 layers.

CNNs work in the same way as typical convolutional networks used for image recognition with the input vector of 496 measurements converted into a 31x16 matrix.

Table 2. The structure of the convolutional neural network

ID	Name	Type	Activations	Learnables
1	imageinput 31x16x1 images with 'zerocenter' normalization	Image Input	31×16×1	-
2	conv_1 96 3x3x1 convolutions with stride [1 1] and padding 'same'	Convolution	31×16×96	Weights 3×3×1×96 Bias 1×1×96
3	batchnorm_1 Batch normalization with 96 channels	Batch Normalization	31×16×96	Offset 1×1×96 Scale 1×1×96
4	relu_1 ReLU	ReLU	31×16×96	-
5	avgpool_1 2x2 average pooling with stride [1 1] and padding [0 0 0 0]	Average Pooling	30×15×96	-
6	conv_2 16 3x3x96 convolutions with stride [1 1] and padding 'same'	Convolution	30×15×16	Weights 3×3×96×16 Bias 1×1×16
7	batchnorm_2 Batch normalization with 16 channels	Batch Normalization	30×15×16	Offset 1×1×16 Scale 1×1×16
8	relu_2 ReLU	ReLU	30×15×16	-
9	avgpool_2 2x2 average pooling with stride [1 1] and padding [0 0 0 0]	Average Pooling	29×14×16	-
10	conv_3 16 3x3x16 convolutions with stride [1 1] and padding 'same'	Convolution	29×14×16	Weights 3×3×16×16 Bias 1×1×16
11	batchnorm_3 Batch normalization with 16 channels	Batch Normalization	29×14×16	Offset 1×1×16 Scale 1×1×16
12	relu_3 ReLU	ReLU	29×14×16	-
13	fc 16 fully connected layer	Fully Connected	1×1×16	Weights 16×6496 Bias 16×1
14	softmax softmax	Softmax	1×1×16	-
15	classoutput crossentropyex	Classification Output	-	-

Fig. 7 shows the course of training the CNN, which task was to classify the pattern for cluster No. 137. It was noted that the cluster No. 137 is a case that makes reconstruction more difficult than other clusters - hence the choice. The mini-batch size for CNN was set to 64. Validation frequency was 30 and number of epochs was 10.

Validation accuracy equals 95.9%. Accuracy is calculated as the ratio of accurately guessed classifiers to all validation cases. Maximum iterations limit was 2960. Training finished because it reached final iteration, not due to the lack of a decrease of the validation error in the next 6 iterations.

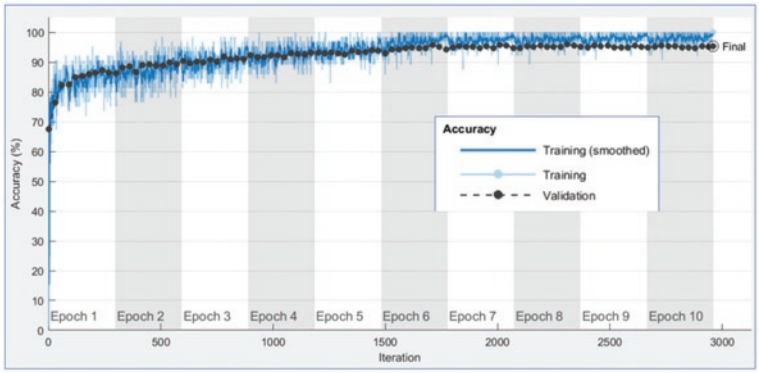


Fig. 7. Training progress through accuracy for cluster No. 137

Fig. 6 corresponds with Fig. 8. It shows training progress through loss indicator for cluster No. 137. The Validation loss is calculated as the ratio of incorrectly guessed classifiers to all validation cases.

By analyzing the shapes of the graphs presented in Fig. 7 and 8 we can conclude that the CNN learning process for the cluster No. 137 proceeded correctly. The initial, large increase in accuracy and the subsequent lack of fluctuations indicate the lack of overfitting. Also, the validation loss at 0.1398 allows us to conclude that CNN has a high generalization capability.

The above results are confirmed in Table 3, where selected iterations, from the CNN learning process for the cluster No. 137, are presented. Base Learning Rate is variable and decreases with subsequent epochs. The algorithm updates the learning speed by reducing the Learning Rate, multiplying it by a specific fractional learn-rate drop factor within a certain number of epochs.

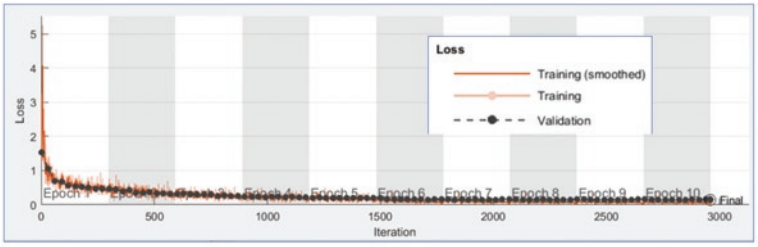


Fig. 8. Training progress through loss indicator for cluster No. 137

Table 3. Training CNN for cluster No. 137 on single GPU

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Validation Accuracy	Mini-batch Loss	Validation Loss	Base Learning Rate
1	1	00:00:00	17.19%	76.20%	2.6260	1.2312	0.0100
1	270	00:00:10	87.50%	87.90%	0.5279	0.3940	0.0100
2	300	00:00:11	89.06%	89.10%	0.4350	0.3766	0.0100
2	570	00:00:21	85.94%	91.50%	0.2722	0.2509	0.0100
3	600	00:00:22	90.63%	91.90%	0.2496	0.2520	0.0100
3	870	00:00:32	95.31%	92.90%	0.2569	0.2173	0.0100
4	900	00:00:33	98.44%	92.80%	0.1025	0.2224	0.0100
4	1170	00:00:43	92.19%	93.80%	0.1765	0.1952	0.0100
5	1200	00:00:44	96.88%	94.30%	0.1214	0.1941	0.0100
5	1470	00:00:54	100.00%	93.80%	0.0361	0.1822	0.0100
6	1500	00:00:55	100.00%	94.10%	0.0669	0.2011	0.0020
6	1770	00:01:05	98.44%	95.40%	0.0371	0.1524	0.0020
7	1800	00:01:06	98.44%	95.60%	0.0862	0.1494	0.0020
7	2070	00:01:16	98.88%	95.90%	0.0855	0.1398	0.0020
8	2350	00:01:27	98.44%	95.90%	0.0971	0.1398	0.0020
10	2960	00:01:51	96.88%	95.90%	0.0855	0.1398	0.0020

3. Results and discussion

3.1. Reconstructions with simulation data

Fig. 9 presents 5 cases of reconstruction, made by applying our algorithm on simulation data for the ultrasound tomography. In order to make an objective assessment of the quality of the reconstruction, the accuracy indicator was used (4):

$$Accuracy=\frac{N_c}{N}\cdot 100\% \tag{4}$$

where: N_c – number of pixels reconstructed correctly, N – total number of pixels.

Individual cases in Fig. 9 were sorted according to the number of inclusions. The first case concerns a single, large inclusions. Its accuracy is the highest and amounts to 99.32%. With the exception of samples 3 and 4, one can see a regularity that the more inclusions, the worse the accuracy of the mappings. This is a typical phenomenon in tomography, which can be observed not only in relation to UST, but in EIT and ECT as well.

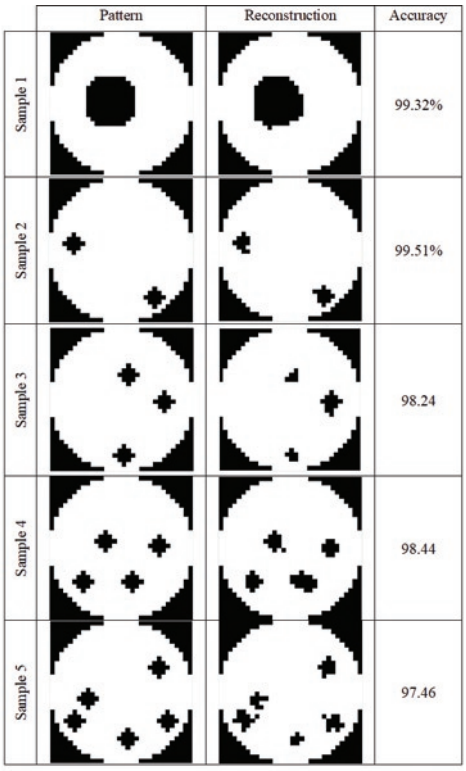


Fig. 9. Simulated reconstructions using MCNN

The general observation of the reconstructed samples shows that in all cases MCNN correctly reflects the size and location of the inclusions. Slightly more difficult are the cases of reconstruction many small objects. Sample 5 in Fig. 9 shows that despite some imperfections, one can still correctly determine both the location and size of the objects hidden in the tank.

Table 4 presents a comparison of the results of a single cluster (No. 137) reconstruction. Because the comparison was for a single cluster, not for the whole image, there was no need to use multiple CNN or multiple ANN.

All six methods listed in Table 4 have 496 measurements at the input, and at the output a classifier generating the numbers $p_i = \{1,2, ..., 16\}$. Cluster 137 is located in the central part of the observation field, which means that it is remote from all transducers. In the case of using CNN, the reconstruction accuracy of the cluster No. 137 is about 2.7% lower than the average accuracy of all reconstructed images listed in Fig. 9, hence the classification for that particular cluster seems to be slightly more difficult than for others. For comparative reconstruction of the cluster No. 137, the following classification technics belonging to the group of machine learning methods were used: Convolutional Neural Network (CNN), Artificial Neural Network (ANN), Linear Support Vector Machine (LSVM), Classification Tree, Medium k-Nearest Neighbor classification (KNN) and Naive Bayes.

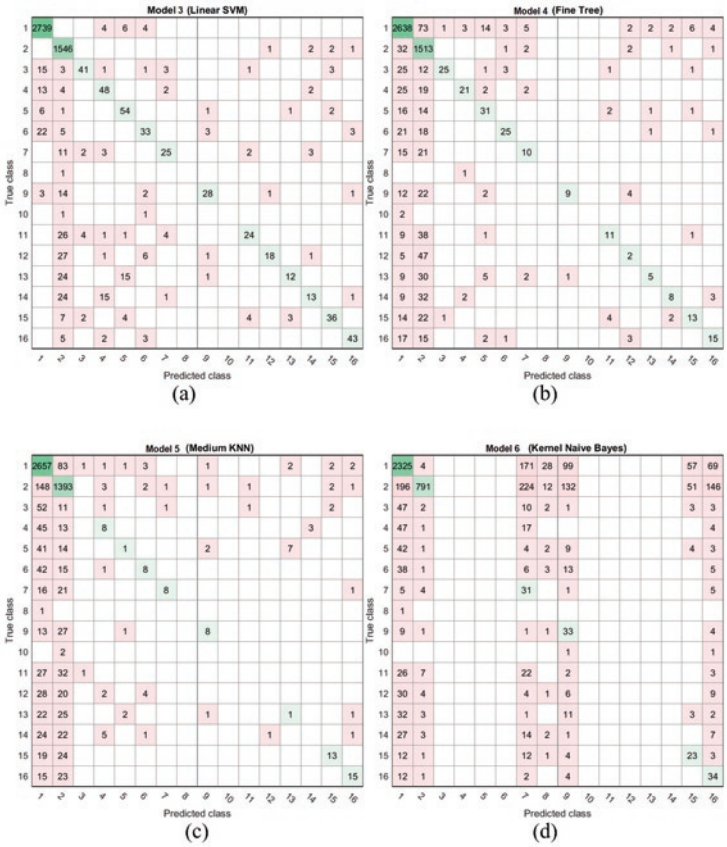


Fig. 11. Confusion matrices for: (a) SVM, (b) Fine Tree, (c) KNN, (d) Naive Bayes

Fig. 10 shows confusion matrices of the cluster No. 137 for two types of neural network: deep CNN (a) and shallow ANN (b) (Table 4, ID 1-2). Fig. 11 shows confusion matrices for four statistical machine learning methods (Table 4, ID 3-6). The comparison of the accuracy indicator showed that the method based on CNN algorithms is the most exact.

Cases correctly reconstructed are marked in green. They are located along the diagonal of the matrix. The number of validation cases used to test accuracy was 4999.

Fig. 10 and 11 should be analyzed in relation to Table 1. The most accurate reconstruction concerned cluster No. 1 (p_i=1 in Table 1).

It is a cluster without inclusions, completely white. This cluster occurs most often at any reconstruction because it depicts the background. Cluster No. 2 is the second most accurate classification (p_i=2 in Table 1). This cluster contains one black pixel in the lower left corner of the 4-pixel area. This cluster can be used for reconstruction of small inclusions, or for imaging the edges of larger inclusions. It can be seen that, for the rest of the cluster, the CNN confusion matrix has

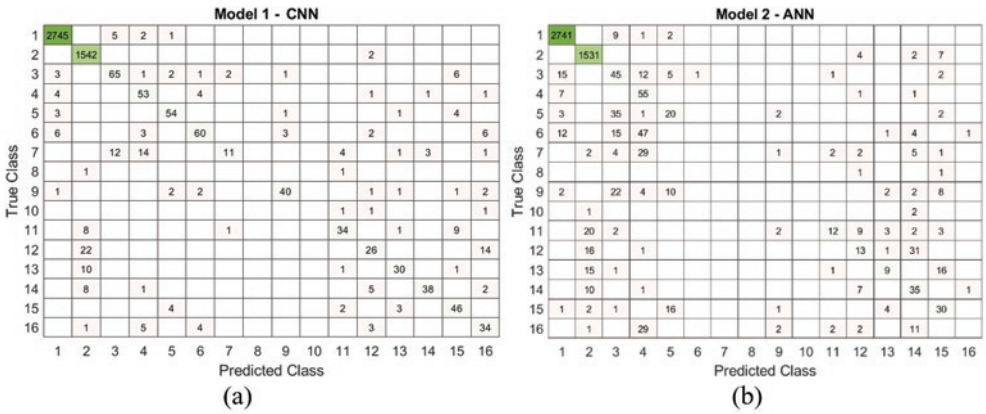


Fig. 10. Confusion matrices for: (a) – deep CNN, (b) – shallow ANN

the most correct hits, as evidenced by the high numbers along the diagonal (Fig. 10a).

Analyzing the results of classical confusion matrices, it should be taken into account that erroneous cases do not have to mean a total mistake in the means of quality of full image. For example two distinct patterns of clusters can differ only by one pixel, thus the whole image can still be reconstructed relatively correctly despite the mistake in classification.

3.2. Reconstructions with real data

Fig. 12 shows the test stand that was used to verify the MCNN algorithm based on real measurements. The measurements obtained from this setup were processed using MCNN trained on the simulation generated data. An adjustable frame was mounted on top of the bucket filled with tap water, allowing variable arrangement of vertical air-filled plastic tubes.

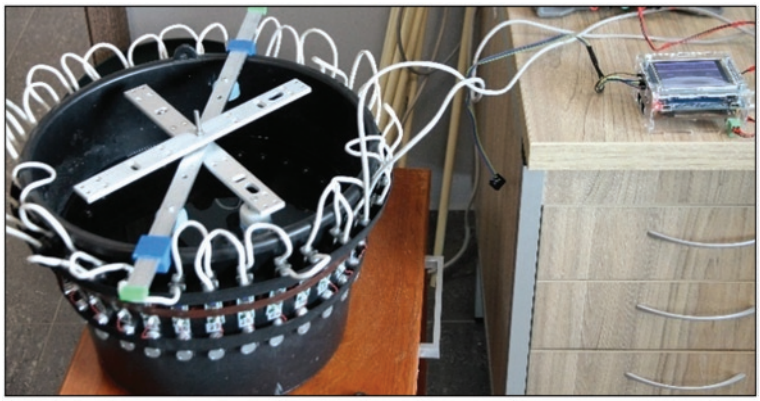


Fig. 12. Stand for collecting real data

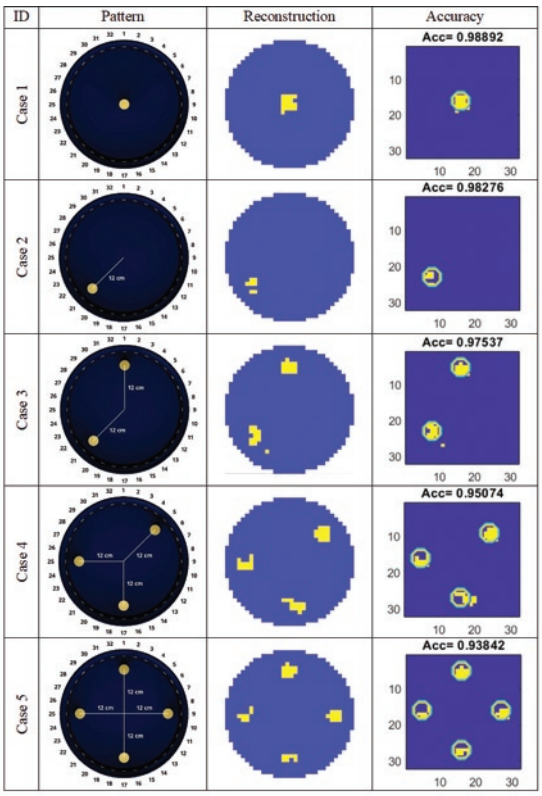


Fig. 13. Real data reconstruction using MCNN

Fig. 13 presents the results of reconstructions based on real measurements. Five cases were tested, which involved a varying number of objects immersed in water. For real reconstructions, reference images were developed based on geometric measurements.

Then the pattern images were applied to a 32×32 pixel matrix. In this way, accuracy was estimated based on relationship (4). Case 1 includes a single centrally located object equidistant from all 32 transducers. The reconstruction of this case is clear. The object has been correctly located, but its shape and dimensions are not reflected with high accuracy. In the tomographic image, the object has a larger diameter and its shape is not rounded.

Case 2 concerned a single object located close to the tank wall. Also in this case the position of the inclusions has been correctly illustrated, but the shape and dimensions are not perfect. In addition, the image of the object is heterogeneous. In fact, these are 2 objects located close to each other. This can be a problem during identifying the real number of inclusions.

Case 3 shows the reconstruction of two objects spaced apart. As in previous cases, the complaints concern the shape and size of reconstructed inclusions. In addition, one of the objects is represented as two separate objects that are close together, one of which is much larger than the other.

Cases 4 and 5 show reconstructions of 3 and 4 inclusions located near the walls of the tank, respectively. Assumptions from the observations are similar to those in previous cases. It is worth noting, however, that although the shapes and dimensions of the objects are not properly reconstructed, their location is determined relatively precisely and the reconstructed image is actually noise free. Based on the conducted experiments, it seems that the UST system works similarly regardless of the number of inclusions and irrespectively from the distance of hidden objects from the transducers.

4. Discussion

Because UST is a method rarely used in process tomography, it is reasonable to ask about the reasons for this. The answer is undoubtedly complex, but the main reasons are two.

The first reason is due to the difficulties in developing the appropriate physical model of sound's interaction [8]. The second reason is the imperfections of data processing algorithms. This work focuses on algorithms, although the UST tomograph used was designed and made by electronic engineers in Netrix SA laboratories. The developed hardware system proved to be good enough to provide sufficient quality data for efficient processing by the deep learning and training algorithm of a multi-convolutional neural network (MCNN) system. The developed method is a new proposal in the field of tomographic algorithms, and although it cannot be said that its use will always outweigh the effectiveness of other known methods, in all tested cases the MCNN algorithm proved to be the most effective.

Tracking the progress of research on algorithms, it can be seen that the unequivocal indication of one method that definitely outweighs the other methods of imaging efficiency in each case is impossible. Reviewing the research results presented by scientists, it can be stated that they depend on specific restrictions related mainly to the details of the research object. It is also known that for the algorithm to work properly, measurements must provide data on sizes and formats closely tailored to given requirements. This fact is the basic barrier hindering the construction of universal tomographs, suitable for a wide range of applications.

In these studies, experiments were conducted using a bucket of water in which plastic pipes were immersed. Simulation algorithms were validated based on the data obtained from measurements at the test stand (Fig. 2b). Thanks to simulation cases, including both precise reference images and their corresponding measurements, it was possible to

compare the effectiveness of several selected algorithms (CNN, ANN, LSVM, Classification Tree, KNN and Naive Bayes). The evaluation criterion was an objective indicator - Accuracy.

In addition to simulation tests, reconstructions of UST images were also made based on real data captured directly from the test stand. Due to the lack of precise reference images, a quantitative assessment of these cases was not possible. On the other hand, it was possible to visually compare the reconstruction with the arrangement diagrams of the inclusions in the individual cases studied. It turned out that despite the expected noise resulting from the presence of many unsteady factors accompanying processes carried out in complex electronic systems, the obtained reconstructions are legible and basically correctly reflect the positions of the inclusions relative to the examined cross-section. This is particularly important taking under consideration the binary specificity of imaging. The binary nature of imaging means that every error is visible in the image as an incorrectly displayed pixel what is a relevant issue due to small resolution of the image.

The limitation of MCNN may be a relatively low reconstruction speed. Although the speed of calculations depends on many factors, the need for a large number of separately working CNNs means that the presented method may not be used in flow systems or systems with high dynamics of change inside the tested tank. Therefore, potential areas of application for UST systems with MCNN algorithms relate to static 2-phase systems, e.g. tanks and reactors with liquid-solid phases, can be limited.

The advantage of MCNN is the ability to properly reconstruct the inclusion position and resistance to noise. Comparing the reconstructed images based on simulation and real data, it can be concluded that the algorithm effectively deals with noise. This is undoubtedly strength of MCNN especially due to the fact that the neural networks was trained on simulation data which are much more easier to collect, in the amount necessary for training process, than the real measurement data.

5. Conclusions

The article presents a new algorithmic method of deep learning, enabling accurate image reconstruction using ultrasound tomography (UST). Known and currently used methods of monitoring tank reactors are still burdened with problems resulting in a relatively low resolution of reconstructed images, hence it was necessary to take up the analyzed subject. The presented tomographic method contributes to improving the diagnostics of technical facilities such as reactors. It enables both early detection of process parameters deviations enabling effective control and detection of hazards resulting in failure. In this context, a higher level of reliability can be achieved by using the developed algorithm.

An important achievement of the research is the noise-resistant algorithm based on multiple convolutional neural networks (MCNN), which, despite being trained on simulation data, effectively reconstructs objects hidden inside the tank, regardless of their shape, quantity, location or dimensions, based on real measurements. The factor that allows achieving good tomographic reconstruction is the training of many neural networks with one cluster output instead of one CNN with multiple outputs. Simulation experiments carried out for selected UST cases have shown that the newly developed MCNN method can be successfully used to generate 2D monochrome images based on the ultrasonic wave time-of-flight measurements.

In order to better verify the quality of the resulting solutions, the CNN was compared with five popular classification methods that could be used interchangeably (Table 4). The comparative analysis of the accuracy indicators for CNN, ANN, LSVM, Classification Tree, KNN and Naive Bayes showed that the newly developed algorithmic method of CNN most accurately reconstructs a single cluster of the image.

As a result of the tests, both based on simulation and real data, it can be concluded that the reconstructions correctly reflect the position of the inclusions. Images obtained from real measurements show slightly too large inclusion diameters. Simulation data images do not have this error. The reason for problems in the correct representation of the inclusion shape is the low resolution of the output image (32x32). It is worth noting that the MCNN algorithm is able to equally well image objects located both in the center of the tank and close to its walls. The algorithm also has the advantage of being able to image both single and multiple inclusions. It is worth noting that the obtained imaging efficiency exceeds 95%.

A way to increase the effectiveness of the algorithm could be to combine EIT methods with UST. This requires the installation of both types of sensors around the reactor under test: electrodes for EIT and transducers for UST. This idea requires some technical problems to be resolved to deploy a large number of different sensors in close proximity.

The topic of the research is up to date, which has been appropriately substantiated with reference to the present state of knowledge. Thanks to the conducted research, it was possible to develop a tomographic algorithm, whose high resistance to noise allows for generating detailed images in real conditions. This was confirmed during special tests. The presented results are of key importance for the development of knowledge and innovation in the field of non-invasive applications for monitoring methods of industrial facilities - especially tank reactors. It is planned to continue research towards the development of a hybrid method covering both the physical (EIT + UST) and algorithmic layers.

References

1. Anders D, Rzasa M. The possibility of composting animal waste products. Environment Protection Engineering 2007;33(2):7-15.
2. Arevalo J, González FA, Ramos-Pollán R, Oliveira JL, Guevara Lopez MA. Representation learning for mammography mass lesion classification with convolutional neural networks. Computer Methods and Programs in Biomedicine 2016;127: 248-257, https://doi.org/10.1016/j.cmpb.2015.12.014.
3. Babout L, Grudzień K, Wiącek J, Niedostatkiewicz M, Karpiński B, Szkodo M. Selection of material for X-ray tomography analysis and DEM simulations: comparison between granular materials of biological and non-biological origins. Granular Matter 2018; 20(3): 38, https://doi.org/10.1007/s10035-018-0809-y.
4. Banasiak R, Wajman R, Jaworski T, Fiderek P, Fidos H, Nowakowski J, Sankowski D. Study on two-phase flow regime visualization and identification using 3D electrical capacitance tomography and fuzzy-logic classification. International Journal of Multiphase Flow 2014; 58: 1-14, https://doi.org/10.1016/j.ijmultiphaseflow.2013.07.003.
5. Bartusek K, Fiala P, Mikulka J. Numerical modeling of magnetic field deformation as related to susceptibility measured with an MR system. Radioengineering 2008;17(4):113-118.
6. Bishop CM. Pattern Recognition and Machine Learning. Springer-Verlag New York 2006.
7. Chen H, Ni D, Qin J, Li S, Yang X, Wang T, Heng PA. Standard Plane Localization in Fetal Ultrasound via Domain Transferred Deep Neural Networks. IEEE Journal of Biomedical and Health Informatics 2015; 19(5): 1627-1636, https://doi.org/10.1109/JBHI.2015.2425041.

8. Du X, Li J, Feng H, Chen S. Image reconstruction of internal defects in wood based on segmented propagation rays of stress waves. Applied Sciences (Switzerland) 2018; 8(10): 1778, https://doi.org/10.3390/app8101778.
9. Goetzke-Pala A, Hoła A, Sadowski Ł. A non-destructive method of the evaluation of the moisture in saline brick walls using artificial neural networks. Archives of Civil and Mechanical Engineering 2018; 18(4): 1729-1742, https://doi.org/10.1016/j.acme.2018.07.004.
10. Gola A, Kłosowski G. Application of Fuzzy Logic and Genetic Algorithms in Automated Works Transport Organization. In: Advances in Intelligent Systems and Computing. Vol 620. Springer, Cham 2018: 29-36, https://doi.org/10.1007/978-3-319-62410-5_4.
11. Khairi MTM, Ibrahim S, Yunus MAM, Faramarzi M, Sean GP, Pusppanathan J, Abid A. Ultrasound computed tomography for material inspection: Principles, design and applications. Measurement 2019; 146: 490-523, https://doi.org/10.1016/j.measurement.2019.06.053.
12. Kłosowski G, Kozłowski E, Gola A. Integer Linear Programming in Optimization of Waste After Cutting in the Furniture Manufacturing. In: Burduk A., Mazurkiewicz D. (Eds) Intelligent Systems in Production Engineering and Maintenance - ISPEM 2017, vol 637, 2018: 260-270, https://doi.org/10.1007/978-3-319-64465-3_26.
13. Kłosowski G, Rymarczyk T, Gola A. Increasing the reliability of flood embankments with neural imaging method. Applied Sciences 2018; 8(9): 1457, https://doi.org/10.3390/app8091457.
14. Kozłowski E., Mazurkiewicz D., Kowalska B., Kowalski D. – Binary Linear Programming as a Decision-Making Aid for Water Intake Operators. In book: Intelligent Systems in Production Engineering and Maintenance – ISPEM 2017, Edition: Advances in Intelligent Systems and Computing vol. 637. Publisher: Springer International Publishing, Editors: Anna Burduk, Dariusz Mazurkiewicz, pp.199-208, DOI: 10.1007/978-3-319-64465-3_20.
15. Krawczyk A, Korzeniewska E. Magnetophosphenes - history and contemporary implications. Przegląd Elektrotechniczny 2018; 1(1): 63-66, https://doi.org/10.15199/48.2018.01.16.
16. Kryszyn J, Smolik W. Toolbox for 3D modelling and image reconstruction in electrical capacitance tomography. Informatics Control Measurement in Economy and Environment Protection 2017; 7(1): 137-145, https://doi.org/10.5604/01.3001.0010.4603.
17. Kryszyn J, Wanta DM, Smolik WT. Gain Adjustment for Signal-to-Noise Ratio Improvement in Electrical Capacitance Tomography System EVT4. IEEE Sensors Journal 2017; 17(24): 8107-8116, https://doi.org/10.1109/JSEN.2017.2744985.
18. Lei J, Liu Q, Wang X. Deep Learning-Based Inversion Method for Imaging Problems in Electrical Capacitance Tomography. IEEE Transactions on Instrumentation and Measurement 2018; 67(9): 2107-2118, https://doi.org/10.1109/TIM.2018.2811228.
19. Li X, Li J, He D, Qu Y. Gear pitting fault diagnosis using raw acoustic emission signal based on deep learning. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(3): 403-410, https://doi.org/10.17531/ein.2019.3.6.
20. Lopato P, Chady T, Sikora R, Gratkowski S, Ziolkowski M. Full wave numerical modelling of terahertz systems for nondestructive evaluation of dielectric structures. COMPEL - The international journal for computation and mathematics in electrical and electronic engineering 2013; 32(3): 736-749, https://doi.org/10.1108/03321641311305719.
21. Ma J, Wu F, Zhu J, Xu D, Kong D. A pre-trained convolutional neural network based method for thyroid nodule diagnosis. Ultrasonics 2017; 73: 221-230, https://doi.org/10.1016/j.ultras.2016.09.011.
22. Majchrowicz M, Kapusta P, Jackowska-Strumiłło L, Sankowski D. Acceleration of image reconstruction process in the electrical capacitance tomography 3D in heterogeneous, multi-GPU system. Informatics Control Measurement in Economy and Environment Protection 2017; 7(1): 37-41, https://doi.org/10.5604/01.3001.0010.4579.
23. Mikulka J. GPU-Accelerated Reconstruction of T2 Maps in Magnetic Resonance Imaging. Measurement Science Review 2015; 15(4): 210-218, https://doi.org/10.1515/msr-2015-0029.
24. Podgórni E, Rząsa M. Investigation of the effects of salinity and temperature on the removal of iron from water by aeration, filtration, and coagulation. Polish Journal of Environmental Studies 2014; 23(6): 2157-2161, https://doi.org/10.15244/pjoes/24927.
25. Psuj G. Multi-Sensor Data Integration Using Deep Learning for Characterization of Defects in Steel Elements. Sensors 2018; 18(2): 292, https://doi.org/10.3390/s18010292.
26. Qayyum A, Anwar SM, Awais M, Majid M. Medical image retrieval using deep convolutional neural network. Neurocomputing 2017; 266: 8-20, https://doi.org/10.1016/j.neucom.2017.05.025.
27. Romanowski A. Big Data-Driven Contextual Processing Methods for Electrical Capacitance Tomography. IEEE Transactions on Industrial Informatics 2019; 15(3): 1609-1618, https://doi.org/10.1109/TII.2018.2855200.
28. Romanowski A. Contextual Processing of Electrical Capacitance Tomography Measurement Data for Temporal Modeling of Pneumatic Conveying Process. In: 2018 Federated Conference on Computer Science and Information Systems (FedCSIS) 2018: 283-286, https://doi.org/10.15439/2018F171.
29. Romanowski A, Łuczak P, Grudzień K. X-ray Imaging Analysis of Silo Flow Parameters Based on Trace Particles Using Targeted Crowdsourcing. Sensors 2019; 19(15): 3317, https://doi.org/10.3390/s19153317.
30. Rymarczyk T. New methods to determine moisture areas by electrical impedance tomography. Kojima F, Kobayashi F, Nakamoto H, eds. International Journal of Applied Electromagnetics and Mechanics 2016; 52(1-2): 79-87, https://doi.org/10.3233/JAE-162071.
31. Rymarczyk T. Using electrical impedance tomography to monitoring flood banks. In: International Journal of Applied Electromagnetics and Mechanics 2014; 45: 489-494, https://doi.org/10.3233/JAE-141868
32. Rymarczyk T, Adamkiewicz P, Duda K, Szumowski J, Sikora J. New electrical tomographic method to determine dampness in historical buildings. Archives of Electrical Engineering 2016; 65(2): 273-283, https://doi.org/10.1515/aee-2016-0019.
33. Rymarczyk T, Kłosowski G, Cieplak T, Kozłowski E, Kania K. Application of a regressive neural network with autoencoder for monochromatic images in ultrasound tomography. In: Institute of Electrical and Electronics Engineers (IEEE) 2019: 156-160, https://doi.org/10.23919/PTZE.2019.8781750.
34. Rząsa MR. A new transducer of double processing for capacitive tomography. Metrology and Measurement Systems 2007; 14(2): 291-305.
35. Rząsa MR, Dobrowolski B. The prototype capacitance tomography sensor with increased sensitivity near the wall. Journal of Energy Science 2010; 1(1):133-145.
36. Sezer A, Basri Sezer H. Convolutional neural network based diagnosis of bone pathologies of proximal humerus. Neurocomputing 2019, https://doi.org/10.1016/j.neucom.2018.11.115.
37. Shahdoosti HR, Rahemi Z. Edge-preserving image denoising using a deep convolutional neural network. Signal Processing 2019; 159: 20-

32, https://doi.org/10.1016/j.sigpro.2019.01.017.
38. Świderski A, Jóźwiak A, Jachimowski R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(2): 292-299, https://doi.org/10.17531/ein.2018.2.16.
39. Szczesny A, Korzeniewska E. Selection of the method for the earthing resistance measurement. Przegląd Elektrotechniczny 2018; 94(12): 178-181.
40. Vališ D, Forbelská M, Vintr Z, Hasilová K, Leuchter J. Platinum thermometer failure estimation based on dynamic linear models. Engineering Failure Analysis 2019; 101: 418-435, https://doi.org/10.1016/j.engfailanal.2019.03.024.
41. Vališ D, Mazurkiewicz D. Application of selected Levy processes for degradation modelling of long range mine belt using real-time data. Archives of Civil and Mechanical Engineering 2018; 18(4): 1430-1440, https://doi.org/10.1016/j.acme.2018.05.006.
42. Wajman R, Fiderek P, Fidos H, Jaworski T, Nowakowski J, Sankowski D, Banasiak R. Metrological evaluation of a 3D electrical capacitance tomography measurement system for two-phase flow fraction determination. Measurement Science and Technology 2013; 24(6): 065302, https://doi.org/10.1088/0957-0233/24/6/065302.
43. Zhang Q, Xiao Y, Dai W, Suo J, Wang C, Shi J, Zheng H. Deep learning based classification of breast tumors with shear-wave elastography. Ultrasonics 2016; 72: 150-157, https://doi.org/10.1016/j.ultras.2016.08.004.
44. Zhao R, Yan R, Chen Z, Mao K, Wang P, Gao RX. Deep learning and its applications to machine health monitoring. Mechanical Systems and Signal Processing 2019; 115: 213-237, https://doi.org/10.1016/j.ymssp.2018.05.050.
45. Ziolkowski M, Gratkowski S, Zywica AR. Analytical and numerical models of the magnetoacoustic tomography with magnetic induction. COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering 2018; 37(2): 538-548, https://doi.org/10.1108/COMPEL-12-2016-0530.

Grzegorz KŁOSOWSKI
Lublin University of Technology
Department of Organization of Enterprise
ul. Nadbystrzycka 38D, 20-618 Lublin, Poland

Tomasz RYMARCZYK
Konrad KANIA
University of Economics and Innovation
ul. Projektowa 4, 20-209 Lublin, Poland
Research and Development Center, Netrix S.A.
ul. Związkowa 26, 20-148 Lublin, Poland

Antoni ŚWIĆ
Lublin University of Technology
Institute of Technological Systems of Information
ul. Nadbystrzycka 38D, 20-618 Lublin, Poland

Tomasz CIEPLAK
Lublin University of Technology
Department of Organization of Enterprise
ul. Nadbystrzycka 38D, 20-618 Lublin, Poland

E-mails: g.klosowski@pollub.pl, tomasz@rymarczyk.com,
konrad.kania@netrix.com.pl, a.swic@pollub.pl, t.cieplak@pollub.pl

Article citation info:
MIELCZAREK M, SŁOWIK M, ANDRZEJCZAK K. The assessment of influence of styrene-butadiene-styrene elastomer's content on the functional properties of asphalt binders. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 148–153, http://dx.doi.org/10.17531/ein.2020.1.17.

Marta MIELCZAREK
Mieczysław SŁOWIK
Karol ANDRZEJCZAK

THE ASSESSMENT OF INFLUENCE OF STYRENE-BUTADIENE-STYRENE ELASTOMER'S CONTENT ON THE FUNCTIONAL PROPERTIES OF ASPHALT BINDERS

OCENA WPŁYWU ZAWARTOŚCI ELASTOMERU STYREN-BUTADIEN-STYREN NA WŁAŚCIWOŚCI FUNKCJONALNE LEPISZCZY ASFALTOWYCH*

This paper discusses the issue of improving the functional properties of road asphalt pavements by modifying bituminous binder with SBS copolymer. The main purpose of the paper is to assess the resistance to permanent deformations and the temperature susceptibility of polymer-modified road asphalt binders, which are most commonly used in the upper layers of road and airport pavements. The bitumens subject to the study originate from various crude oil deposits (Russian and Venezuelan). They were modified in laboratory conditions with a concentrated additive with the known content of the SBS copolymer of 9%. The result was a asphalt binder containing the known percentage of the SBS copolymer of 1.5%, 3.0%, 4.5% and 6%. The rheological properties of the tested bitumens were determined by use of a dynamic shear rheometer (DSR), and with the application of the sinusoidal variable load, in the broad test temperature spectrum (from 40°C to 100°C). The analysis of the values of the dynamic shear modulus |G| of all the studied bitumens shows that the increase in the content of SBS copolymer in the tested binder increases the value of |G*|, which may result in higher resistance to permanent deformations of road pavements caused by repeated traffic loads, especially in the case of pavements operated at high temperatures. The asphalt mixtures resistance to rutting is one of the basic parameters related to road pavement service-life, affecting both the safety and driving comfort of users.*

***Keywords**: dynamic shear rheometer (DSR), rutting factor, copolymer SBS, bitumen, complex shear modulus.*

Tematyka pracy związana jest z zagadnieniem polepszenia właściwości funkcjonalnych drogowych nawierzchni asfaltowych poprzez modyfikację lepiszcza asfaltowego kopolimerem SBS. Głównym celem pracy jest ocena odporności na odkształcenia trwałe oraz wrażliwości na zmiany temperatury asfaltów drogowych modyfikowanych polimerami, które są najczęściej używane w wierzchnich warstwach konstrukcji nawierzchni drogowych i lotniskowych. Przedmiotem badań były asfalty pochodzące z różnych złóż ropy naftowej (rosyjskiej i wenezuelskiej). Asfalty te poddano modyfikacji w warunkach laboratoryjnych z dodatkiem koncentratu o znanej zawartości kopolimeru SBS równej 9%. Otrzymano w ten sposób lepiszcza asfaltowe o znanej zawartości kopolimeru SBS równej 1,5%; 3,0%; 4,5% oraz 6%. Właściwości reologiczne badanych asfaltów oznaczono z użyciem reometru dynamicznego ścinania DSR stosując w testach obciążenie sinusoidalnie zmienne, w szerokim zakresie temperatury pomiarowej (od 40°C do 100°C). Analizując wartości dynamicznego modułu ścinania |G| wszystkich badanych asfaltów można stwierdzić, iż wzrost zawartości kopolimeru SBS w badanym lepiszczu zwiększa wartość |G*|, co może skutkować większą odpornością na odkształcenia trwałe nawierzchni drogowej spowodowane wielokrotnie powtarzającymi się obciążeniami ruchem pojazdów, w szczególności w przypadku nawierzchni eksploatowanej w wysokiej temperaturze Odporność mieszanek mineralno-asfaltowych (MMA) na powstawanie kolein jest jednym z podstawowych parametrów związanych z eksploatacją nawierzchni drogowych, wpływając zarówno na bezpieczeństwo, jak i komfort jazdy użytkowników.*

***Słowa kluczowe**: reometr dynamicznego ścinania, wskaźnik odkształcalności, kopolimer SBS, asfalt, dynamiczny moduł ścinania.*

1. Introduction

Nowadays, road pavements are subject to ever increasing traffic loads [4]. Taking into account both the construction, as well as maintenance costs [6], it is reasonable to search for solutions to optimise, for example, the composition of materials used for pavement structure. The analysis of the test results presented in [1, 11, 12] shows that one of the key factors increasing pavement rutting is the composition of the asphalt mixture, especially the applied asphalt binder. Therefore, there are studies striving to achieve the best rheological properties for the applied bitumens processed in the crude oil distillation. These properties can be improved by introducing various modifiers to the binder structure, i.e. polymers [5], crumb rubber from car tyres [6] or natural asphalts [7]. Many scientific papers have analysed the effects of bitumen modification with the most commonly used polymers, including plastomers (e.g. polyethylene, polypropylene, ethylene vinyl acetate [17]), thermoplastic elastomers: SBS (styrene-butadiene-styrene)) [1, 16]; SIS (styrene-isoprene-styrene) or mixed modifiers consisting of various polymers [2, 12]. Binders modified with polymers show an improvement in rheological properties com-

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl***

pared to unmodified bitumens [20]. The polymer most widely used in road construction is the block copolymer SBS, which, when added to hot bitumen, increases its volume several times in relation to its initial value [21]. When the concentration of the polymer in a modified bitumen is about 6%, the polymer becomes the dispersion phase and forms a continuous network in the structure of the bitumen. In case of a lower SBS concentration, the polymer network does not have to be continuous. Therefore, it is vital, both for the technical and economical reasons, to set the boundary content of the SBS copolymer in the asphalt binder in which the polymer network can form. Elastomer modified bitumens at operating temperatures are characterised by immediate elasticity (elastic deformation) and delayed (creep) elasticity [19, 21]. The papers [1, 12, 21] present modification methods and benefits of using SBS copolymer to modify binder, i.e. higher softening point, decrease in temperature susceptibility (expansion of the temperature-related viscoelasticity range), increase in cohesion at low temperature, significant improvement to elastic properties (observed e.g. in the elastic recovery test). Improving the rheological properties of binders results in better characteristics of the asphalt mixtures treated with the modified binder [13], that is increase in the resistance to permanent deformation and thermally induced cracking.

Airey in paper [1] analysed bitumens originating from two crude oil deposits (Russian and Venezuelan). The bitumens were modified with SBS copolymer, the concentration of the polymer in the modified bitumen amounting to 3%, 5%, and 7%, respectively. He noticed a significant influence of the polymer on rheological properties of the modified bitumen, i.e. increase of dynamic shear modulus and a higher proportion of the elastic part in the bitumen, especially at high temperature. He has also demonstrated the compatibility problem of the bitumen-polymer system. Airey [1] proved that bitumens with paraffin wax (of the Russian origin), due to higher content of the aromatic group, are better at binding the polymer in the modified bitumen structure. Behnood and Olek [6] performed a comparative analysis of three types of modifiers: SBS copolymer, crumb rubber and polyphosphoric acid. The low-temperature properties were studied with a bending beam rheometer (BBR), while the rheological properties at high temperatures were determined with a dynamic shear rheometer (DSR). At high temperatures, they observed an increase in the value of dynamic shear modulus for all the bitumens tested. Bitumen used in the production of asphalt mixtures, in road pavements, is exposed to ageing processes both during storage, production of asphalt mixtures, transport, paving and operation of the pavement [22]. Ageing phenomenon occurring during the production of asphalt mixtures and paving process [24] is considered to be the most unfavourable danger due to high temperatures. During short-term (technological) ageing occurring in the production and construction process of the asphalt pavement, bitumen is exposed to high temperatures (140 – 200˚C) and oxygen in the air. Airey [1] and Sarnowski [19] have demonstrated the problem concerning the ageing of modified bitumen. Modified binders show improved rheological properties in the wide viscoelastic range. The authors observed that once modified bitumen has undergone the process of ageing, it shows a higher proportion of viscous part in relation to the elastic part, which may be caused by partial degradation of the polymer at high temperature, which occurs during the technological processes of coating the aggregate with bitumen, and the transport, paving and compaction of asphalt mixtures [1]. The next stage of ageing takes place during pavement operation (long-term ageing). In this case, binders are exposed to temperatures of up to 60˚C in the summer season, with simultaneous exposure to oxygen, sunlight, water and chemicals [16,22]. Bai [6] performed a study on the influence of short-term and long-term ageing on rheological properties of asphalt binder modified with SBS copolymer. The tests were run on three bitumens with different polymer content, that is 3%, 6% and 9%. Based on the tests using the dynamic shear rheometer, Fraass Breaking Point apparatus and penetrometer, they proved the negative influence of ageing on low-temperature properties of polymer modified bitumen.

The needs for maintenance and repair of the road network are very high. The systematically increasing traffic loads have adverse effect on the condition of the pavements, quickening their degradation [15]. Examining pavement condition encompasses such important means as pavement diagnostics and functional properties testing, i.e. longitudinal evenness, transverse evenness (ruts) [8], friction coefficient, load capacity, etc. [14, 15, 23]. Basic types of asphalt pavement destruction include: rutting, fatigue cracking, and low-temperature cracking [10, 13]. The formation of permanent deformations (ruts) in road pavements is influenced by many factors [13, 14], including the applied aggregate, binder, asphalt mixture, climate conditions, traffic load [9], and the applied pavement structure.

The purpose of the presented research work was to analyse the functional properties of binders modified with SBS copolymer, with special consideration of its temperature susceptibility, as the type of the applied binder is one of the key factors affecting the resistance to permanent deformation (ruts) in asphalt pavements. The resistance to rutting of asphalt mixtures is one of the basic preconditions for proper road pavement operation, affecting both safety and comfort of road users. This paper puts forth an original achievement of applying temperature susceptibility analysis of the modified binders tested at a wide range of temperatures.

2. Properties of the studied binders

Modification of bitumens with SBS copolymer is usually performed in the refineries, and rarely in the road construction companies' installations. Polymer modified bitumen used for the production of asphalt mixture can be obtained by purchasing a ready-made modified binder from a refinery, producing bitumen modified in a special technological installation, or purchasing bitumen with a known mass content of SBS copolymer, e.g. 9%, and mixing it in appropriate proportions with petroleum bitumen (road bitumen) [21].

The research was carried out using 50/70 penetration grade bitumens of similar hardness, expressed as penetration value determined at 25˚C (Table 1), produced from crude oil from Venezuela and Russia. In the conducted studies, bitumens were combined with a concentrate of bitumen modified with SBS copolymer (block copolymer with linear structure) with 9% polymer content by composing in the proportions 5:1, 2:1, 1:1, and 1:2 to obtain asphalt binders with SBS copolymer content: 1.5%; 3.0%; 4.5% and 6.0% (in relation to the

Table 1. Basic properties of the tested asphalt binders

Properties / Tested material	$T_{R\&B}$ [°C]	Pen_{25} [mm/10]
V50/70	47.4 ± 0.2	66.0 ± 0.8
V1.5%SBS	47.9 ± 0.3	69.5 ± 0.4
V3%SBS	52.8 ± 1.3	71.3 ± 0.8
V4.5%SBS	74.5 ± 2.0	66.4 ± 0.6
V6%SBS	87.9 ± 1.3	66.3 ± 0.8
R50/70	47.8 ± 0.4	69.3 ± 0.3
R1.5%SBS	48.8 ± 0.3	70.3 ± 0.3
R3%SBS	49.5 ± 0.3	71.3 ± 0.4
R4.5%SBS	77.0 ± 1.1	69.5 ± 0.5
R6%SBS	83.8 ± 06	71.9 ± 1.0
K9%SBS	100.3 ± 1.4	74.3 ± 1.0

where: $T_{R\&B}$ - softening point determined acc. to PN-EN 1427:2015-08, Pen_{25} - penetration in 25°C determined acc. to PN-EN1426:2015-08

weight of the obtained modified bitumen), respectively. The specimens of the asphalt binders have been marked in the paper by the bitumen's origin, and then the percentage content of SBS polymer, e.g.:

- R6%SBS – means bitumen produced from Russian crude oil with a content of 6.0% of SBS copolymer,
- V50/70 – means 50/70 bitumen from Venezuelan crude oil, containing no SBS copolymer,
- K9%SBS – means modified bitumen concentrate containing 9.0% SBS.

Bitumen was analysed both in its initial state and after the process of technological (short-term) ageing simulated with the RTFOT (Rolling Thin Film Oven Test) method according to PN-EN 12607-1:2014.

The analysis of the results presented in Table 1 shows that the asphalt binders were selected so as to obtain bitumens of similar hardness, expressed through penetration determined at 25°C (Pen_{25} was obtained in the range 66.0 mm/10 to 74.3 mm/10). Therefore, all modified binders subjected to the study can be classified as 45/80 modified bitumens, which are available on the Polish market, although they have different percentages of SBS copolymer content.

3. Purpose and methodology of the studies

The main aim of the study was to evaluate the resistance to permanent deformation and temperature susceptibility of road bitumens modified with SBS copolymer on the basis of tests carried out with dynamic shear rheometer (DSR) of the Physica MCR 101 type. Permanent deformation and sensitivity to temperature changes in the Polish climate zone are of key importance in the use of asphalt pavements.

The tests were performed in compliance with the norm: PN-EN 14770:2012 "Determination of complex shear modulus and Phase Angle - Dynamic Shear Rheometer". They employed two methods involving kinematic (sinusoidal) coercion:

a) at different angular frequency ranges from 100 rad/s to 0.1 rad/s, and a constant test temperature of 60°C±0.01˚C,
b) with a deflection angle amplitude of 10 mrad and variable temperature, i.e. from 100°C to 40°C, and a temperature decrease of 1°C every 1 minute. The test procedure involved an assumption of a constant angular frequency value equal to 10 rad/s.

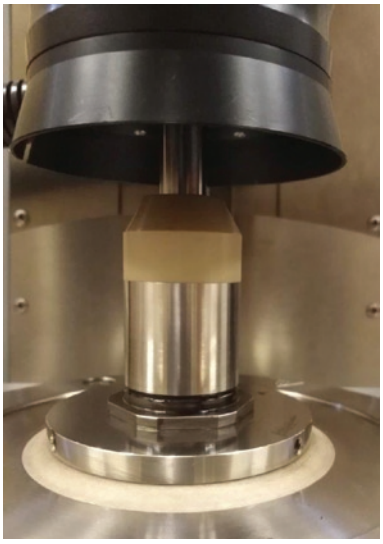


Fig. 1. A side-view of the tested bitumen sample

Fig. 2. Dynamic Shear Rheometer (DSR) of the Physica MCR 101 type

A sample of the asphalt binder was placed between two circular parallel plates with a diameter of Ø25 mm, maintaining the required gap height of 1 mm (Fig. 1).

According to the American Superpave specification, the susceptibility of bitumen binders to permanent deformation in road pavements is determined by the rutting factor, expressed as a value of |G*|/sin δ. In this paper, the above mentioned coefficient was determined for bitumens subjected to the RTFOT short-term ageing method, as well as for bitumens not subjected to it.

The analysis also included the values of the Shear Modulus Index (SMI), a measure of the temperature susceptibility of the binders studied, calculated according to the formula [20]:

$$SMI_{T_2/T_1} = \left| \frac{\log\log\left|G^*_{T_1}\right| - \log\log\left|G^*_{T_2}\right|}{\log\left(T_1 + 273{,}15\right) - \log\left(T_2 + 273{,}15\right)} \right| \quad (1)$$

where:

SMI - Shear Modulus Index

$\left|G^*_{T_1}\right|;\left|G^*_{T_2}\right|$ - dynamic shear modulus at temperature T_1, T_2, respectively, [Pa]

$T_1;T_2$ – extreme temperatures of the measurements taken using DSR, where $T_1 > T_2$, [°C]

It was assumed in this paper that $T_1 = 100°C; T_2 = 40°C$.

4. Analysis of test results

Figure 3 shows a graph of the relation of the angular frequency to the dynamic shear modulus of the tested bitumen (range from 0.1 rad/s to 100 rad/s). For all tested bitumens, the dynamic shear modulus values increases with the growth of angle frequency. The highest value of |G*| at an angular frequency of 0.1 rad/s has been observed for the bitumen with the content of SBS copolymer equal to 9%, while the lowest value has been found for the R50/70 bitumen. With an angular frequency of 100 rad/s, the dynamic shear module achieves values close to from 23680Pa for K9%SBS to 41010Pa for V50/70 bitumen.

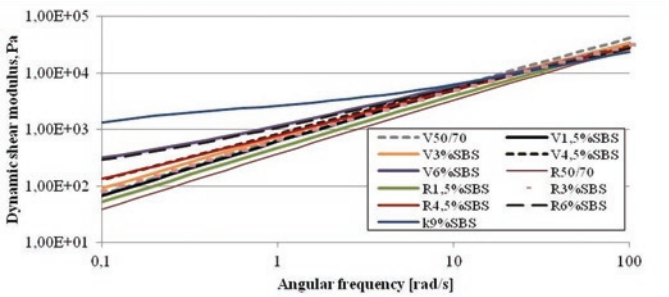


Fig. 3. Graph showing the relation of the angular frequency to the dynamic shear modulus of the bitumens tested at 60˚C

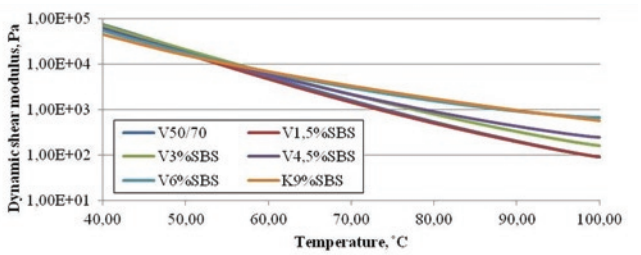


Fig. 4. A graph showing temperature dependence of |G| dynamic shear modulus for unaged Venezuelan binders, at a constant angular frequency value of 10 rad/s*

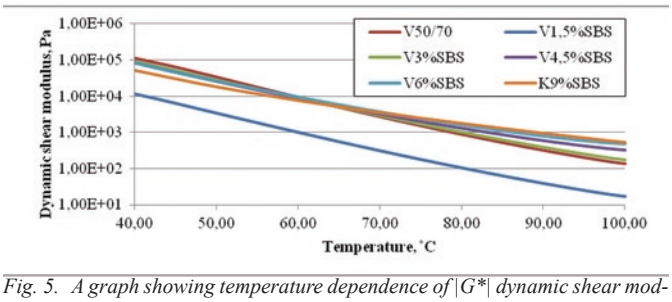


Fig. 5. A graph showing temperature dependence of |G| dynamic shear modulus for Venezuelan binders subjected to RTFOT ageing, at a constant angular frequency value of 10 rad/s*

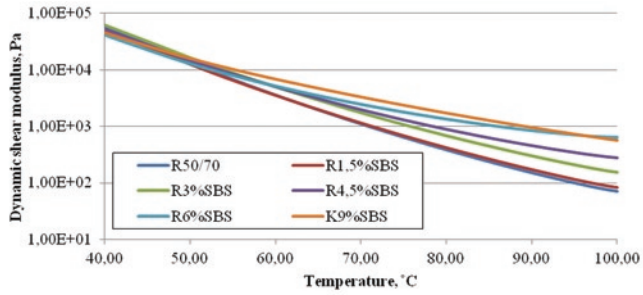


Fig. 6. A graph showing temperature dependence of |G| dynamic shear modulus for unaged Russian binders, at a constant angular frequency value of 10 rad/s*

Asphalt binders are materials with viscoelastic properties. The analysis of the values of the phase angle δ can be used to evaluate changes in the rheological properties of bitumen across the entire temperature spectrum during the production process, paving and operation of the asphalt pavement. Viscous materials are characterised by the damping factor tgδ → ∞ (δ = 90°), whereas in the case of elastic materials tgδ = 0 (δ = 0°); viscoelastic materials have a phase angle values of 0° < δ < 90°. In Figures 4-7 there are graphs showing temperature dependence of |G*| dynamic shear modulus of the studied bitumen binders of Venezuelan and Russian origin, both submitted to, and not having been subject to RTFOT ageing. The value of the |G*| dynamic shear modulus decreases with the increase of the test temperature for all analysed bitumens.

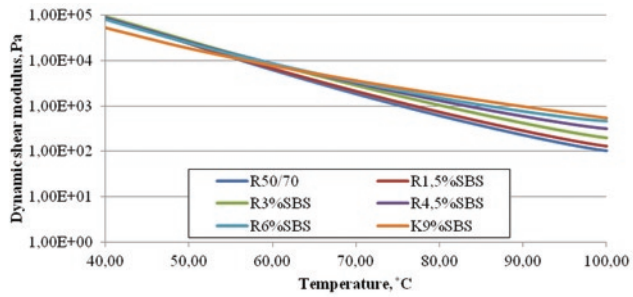


Fig. 7. A graph showing temperature dependence of |G| dynamic shear modulus for Russian binders subjected to RTFOT ageing, at a constant angular frequency of 10 rad/s*

In view of the high summer temperatures, the elastic component is especially important, which is associated with low tgδ values. It was observed that in the case of unmodified bitumens (R and V) and bitumens with a low content of SBS copolymer (up to 3%), there is a regularity that the higher the values of the dynamic shear modulus, the lower the values of the phase angle δ, as shown in the Black diagram (Fig.8.). Based on the graphs showing the relation of the dynamic shear modulus to the phase angle, called Black's graphs, it is possible to perform an analysis of two basic parameters determined in the dynamic shear rheometer [1,3,20]. In the case of asphalt binders with 6% and 9% SBS copolymer content, small phase angle values are obtained, both at very low and high |G*| values. The highest variation in phase angle was observed for reference 50/70 penetration grade bitumens and low-modified binders (with up to 3% copolymer content). Their values at high temperatures are close to 90°, so it can be assumed that these binders at the high temperature range have properties similar to those of a viscous liquid. The increase in the content of the SBS copolymer in the bitumen makes the variation in δ values ever smaller. Above 70°C, the δ values are reduced for bitumens with a copolymer content of 4.5%, 6%, and 9% SBS concentrate, which shows the beneficial effect of applying the polymer for modification, since modified binders have a higher share of the elastic part at high temperatures, which may indicate greater resistance to permanent deformation.

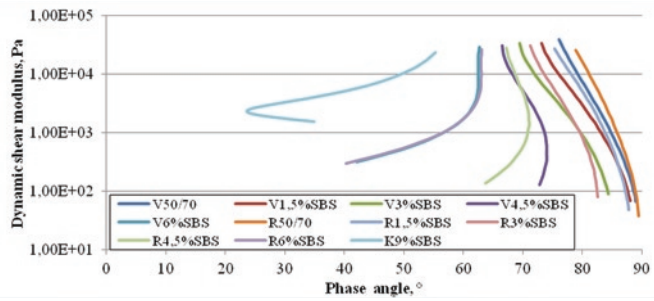


Fig. 8. Black's graph showing the relation of the phase angle to the dynamic shear modulus in the tested asphalt binders

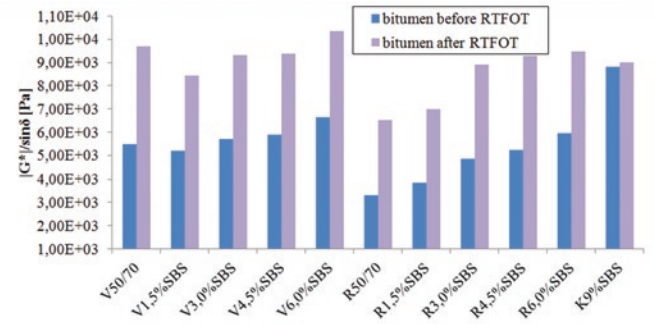


Fig.9. Values of |G|/sinδ rutting factor of the asphalt binders tested at 60°C.*

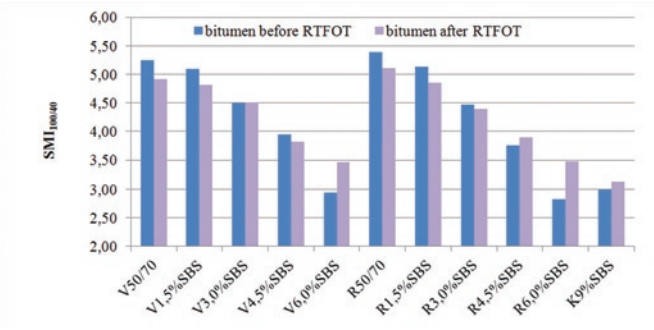


Fig. 10. Shear Modulus Index of the tested bitumen binders

As an effect of conducted tests, the rutting factor was determined (defined as the ratio of the |G*| dynamic shear modulus to the sinus of the phase angle δ (|G*|sinδ)) at the temperature of 60°C, assumed as the extreme temperature occurring in asphalt pavements in Poland (in which the rutting test for asphalt mixtures according to PN-EN 12697-22:2008 is also performed).

The Superpave specification indicates a relation between the resistance to permanent deformation in bitumen pavements and the

properties of the tested binders determined by the DSR method, with the following requirements:

$|G^*|/\sin\delta \geq 1{,}0$ kPa – for unaged bitumen not submitted to ageing

$|G^*|/\sin\delta \geq 2{,}2$ kPa – for bitumen subjected to short-term ageing simulated with the RTFOT method.

The results shown in Figure 9 show that all the binders tested meet the above requirements of the Superpave specifications. A higher value of the rutting factor of the bitumen characterised with the resistance to permanent deformation of the asphalt pavement is obtained by a higher value of the dynamic shear modulus |G*| and a lower value of the phase angle δ.

The lowest value of the rutting factor was observed for unmodified bitumens of Russian origin, both before and after ageing with RTFOT method. Special attention should be paid to the value of $|G^*|/\sin\delta$ of the concentrate containing 9% of SBS copolymer content, since the difference in the rutting factor before and after ageing is only 0.2 kPa; which may indicate that the effect of ageing on the values of the rutting factor is low.

The measure of temperature susceptibility is expressed with the PI penetration index. It can be calculated on the basis of penetration results determined at two various temperatures or by means of an indirect method using values of penetration determined at 25°C and softening point. The above methods make it possible to estimate the penetration index of unmodified bitumens. However, in the case of bitumens modified with elastomers, the results obtained with each method may significantly differ [20]. In this study, the temperature susceptibility was determined on the basis of formula (1). It can be stated that the assumptions for penetration at softening point (800 mm/10) and Fraass breaking point (1.25 mm/10) for bitumen modified with elastomers are not correct [20]. The analysis of the obtained SMI values showed a remarkable influence of the copolymer content on the reduction of bitumen sensitivity to changes in stiffness to temperature. Bitumen R50/70 proved to be the most sensitive to changes in properties due to temperature changes. The lowest SMI value at the temperature range 100°C – 40°C was achieved for the bitumen with SBS content equal to 6% in the unaged state, while upon RTFOT ageing, the SMI value increased for this group of bitumens (both of Venezuelan and Russian origin), which may indicate partial polymer degradation under the influence of high temperature (163°C) and oxygen.

5. Conclusions

By comparing the values of the dynamic shear modulus of the tested bitumens it can be stated that with the increase in the content of SBS copolymer the value of |G*| increases, which may indicate a higher resistance of asphalt pavements made with SBS copolymer modified bitumens to deformations caused by repeated shear stress (which illustrates repeated load cycles caused by traffic in real conditions).

With the increase in the content of SBS copolymer in the tested bitumens, the value of phase angle decreases, which results in the improvement of elastic properties of the binders.

SMI (Shear Modulus Index) analysis showed a considerable effect of styrene-butadiene-styrene copolymer (SBS) content on the reduction of the sensitivity of asphalt binders to changes in stiffness at variable temperatures. Reducing the susceptibility to temperature changes illustrates the significantly favourable effect of using SBS copolymer as a modifier of asphalt binders.

The use of binders modified with SBS copolymer in asphalt mixtures improves the functional properties of flexible pavements (which is confirmed by the values of the rutting factor and SMI), thus, improving the operational parameters of road pavements and their durability.

References

1. Airey G. Rheological properties of styrene butadiene styrene polymer modified road bitumens. Fuel 2003; 82: 1709-1719, https://doi.org/10.1016/S0016-2361(03)00146-7.
2. Ahmedzade P. The investigation and comparison effects of SBS and SBS with new reactive thermopolymer on the rheological properties of bitumen. Construction and Building Materials 2013; 38: 285-291, https://doi.org/10.1016/j.conbuildmat.2012.07.090.
3. Andriescu A, Hesp S.A.M. Time-temperature superposition in rheology and ductile failure of asphalt binders. International Journal of Pavement Engineering 2009; 10(4): 229-240, https://doi.org/10.1080/10298430802169440.
4. Andrzejczak K. Zmiany wzrostu wskaźnika nasycenia samochodami osobowymi. Wiadomości Statystyczne 2012; 11: 22-33.
5. Bai M. Investigation of low-temperature properties of recycling of aged SBS modified asphalt . Construction and Building Materials 2017; 150: 766-773, https://doi.org/10.1016/j.conbuildmat.2017.05.206.
6. Behnood A, Olek J. Rheological properties of asphalt binders modifies with styrene-butadiene-styrene, ground tire rubber (GTR), or polyphosphoric acid (PPA). Construction and Building Materials 2017; 151: 464 - 478, https://doi.org/10.1016/j.conbuildmat.2017.06.115.
7. Bilski M, Słowik M. Impact of aging on Gilsonite and Trinidad Epuré modified binders resistance to cracking. Bituminous Mixtures and Pavements VII 2019; I: 65-70, https://doi.org/10.1201/9781351063265-11.
8. Bogdański B, Słowik M. Analiza porównawcza odporności na koleinowanie mieszanek mineralno-asfaltowych z uwzględnieniem kryteriów oceny wg metody francuskiej (LCPC) i brytyjskiej (BS). Międzynarodowa Konferencja Naukowo-Techniczna "Nowoczesne technologie w budownictwie drogowym" Poznań 2001: 300-309.
9. Cheng Z, Remenyte-Prescott R. Two probabilistic life-cycle maintenance models for the deteriorating pavement. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(3): 394-404, https://doi.org/10.17531/ein.2018.3.7.
10. D'Angelo J, Kluttz R, Dongré R, Stephens K, Zanzotto L. Revision of the Superpave High Temperature Binder Specification: The Multiple Stress Creep Recovery Test. Journal of the Association of Asphalt Paving Technologists 2007; 76: 123-157.
11. Gajewski M, Sybilski D, Bańkowski W. The influence of binder rheological properties on asphalt mixture permanent deformation. The Baltic Journal of Road and Bridge Engineering 2015; 10(1): 54-60, https://doi.org/10.3846/bjrbe.2015.07.
12. Gajewski M, Sybilski D, Bańkowski W, Wróbel A, Mirski K. Ocena odporności na deformacje trwałe mieszanek mineralno-asfaltowych na podstawie zaproponowanego parametru funkcjonalnego lepiszcza. Część 1. Badania lepiszczy. Drogownictwo 2009; 9: 279-283.
13. Gajewski M, Wróbel A, Jemioło S, Sybilski D. Wpływ właściwości reologicznych lepiszcza na koleinowanie MMA. XIV International Conference Computer Systems Aided Science 2010: 857-866.
14. Judycki J, Jaskuła P. Diagnostyka i modernizacja konstrukcji nawierzchni drogowych. Diagnostyka, monitoring i modernizacja eksploatowanych obiektów budowlanych 2010; 56: 233-252.
15. Levulyté L, Žuraulis V, Sokolovskij E. The research of dynamic characteristics of vehicle driving over road roughness. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(4): 518-525.

16. Lu X, Isacsson U. Effect of ageing on bitumen chemistry and rheology. Construction and Building Materials 2002; 16: 15-22, https://doi.org/10.1016/S0950-0618(01)00033-2.
17. Merijs-Meri R, Abele A, Zicans J, Haritinovs V. Development of polyolefine elastomer modified bitumen and characterization of its rheological and structural properties. Bituminous Mixtures and Pavements VII 2019; I: 51- 57, https://doi.org/10.1201/9781351063265-9.
18. Radziszewski P. Wpływ modyfikacji elastomerem SBS na właściwości reologiczne lepiszczy asfaltowych. Polimery 2008: 559-563, https://doi.org/10.14314/polimery.2008.559.
19. Sarnowski M. Rheological properties of road bitumen binders modifies with SBS polymer and polyphophosphoric acid. Roads and Bridges - Drogi i Mosty 2015; 1: 47-65.
20. Słowik M. Thermorheological Properties Of Styrene-Butadiene-Styrene (SBS) Copolymer Modified Road Bitumen. Procedia Engineering 2017; 208:145-150, https://doi.org/10.1016/j.proeng.2017.11.032.
21. Słowik M. Wybrane zagadnienia lepkosprężystości drogowych asfaltów modyfikowanych zawierających elastomer SBS. Rozprawy series Publishing House of Poznan University of Technology 2013; 508.
22. Słowik M, Adamczak P. Ocena wpływu starzenia krótkoterminowego na właściwości asfaltów drogowych modyfikowanych elastomerem SBS. Roads and Bridges - Drogi i mosty 2007; 1: 41-58
23. Surblys V, Žuraulis V, Sokolovskij E. Estimation of road roughness from data of on-vehicle mounted sensors. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19(3): 396-374, https://doi.org/10.17531/ein.2017.3.7.
24. Zicans J, Ivanova T, Merijs-MEri R, Berzina R, Haritonovs V. Aging behavior of bitumen and elastomer modified bitumen. Bituminous Mixtures and Pavements VII 2019; I: 46-50, https://doi.org/10.1201/9781351063265-8.

Marta MIELCZAREK
Mieczysław SŁOWIK
Institute of Civil Engineering
Poznan University of Technology,
ul. Piotrowo 5, 60-965 Poznań, Poland

Karol ANDRZEJCZAK
Institute of Mathematics
Poznan University of Technology,
ul. Piotrowo 3A, 60-965 Poznań, Poland

Emails: marta.mielczarek@put.poznan.pl,
mieczyslaw.slowik@put.poznan.pl, karol.andrzejczak@put.poznan.pl

Article citation info:
SZUMSKA E, JURECKI R, PAWEŁCZYK M. Evaluation of the use of hybrid electric powertrain system in urban traffic conditions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 154–160, http://dx.doi.org/10.17531/jein.2020.1.18.

Emilia SZUMSKA
Rafał JURECKI
Marek PAWEŁCZYK

EVALUATION OF THE USE OF HYBRID ELECTRIC POWERTRAIN SYSTEM IN URBAN TRAFFIC CONDITIONS

OCENA ZASTOSOWANIA NAPĘDÓW HYBRYDOWYCH W WARUNKACH RUCHU MIEJSKIEGO

The conditions of use of the vehicle significantly affect the performance results. Traffic conditions in a specific city directly affect the consumption of energy, fuel and emissions of harmful compounds in exhaust fumes. Conduction of the measurements of a vehicle's performance parameters in operating conditions is very troublesome and is often not possible to realize. An alternative is to use the simulation programs. Vehicle simulation programs offer options related to vehicle models or drive unit components and allow development of new models. Based on the results of simulation testing, it is possible to analyse the level of fuel and energy consumption as well as emissions of harmful compounds in exhaust gases and the operating effectiveness of the drive system in the speed profile. The paper presents the evaluation of the effectiveness of using hybrid electric drive system in passenger cars in medium-sized city traffic conditions using the Kielce example. The simulation tests were based on the speed profiles recorded during real-world test drives in various times of the day. The simulation results were used to conduct an analysis of fuel consumption and pollutant emissions recorded by conventional and hybrid vehicles.

***Keywords**: hybrid electric vehicles, real-world conditions, fuel economy, air pollutants.*

Warunki użytkowania pojazdu mają znaczący wpływ na parametry eksploatacyjne pojazdu. Warunki ruchu w określonym mieście bezpośrednio wpływają na zużycie energii, paliwa i poziom emisji szkodliwych związków zawartych w spalinach. Przeprowadzenie pomiarów parametrów eksploatacyjnych pojazdu w warunkach rzeczywistych jest kłopotliwe i często niemożliwe do zrealizowania. Alternatywą jest wykorzystanie symulacji komputerowych. Programy do symulacji pojazdów oferują, między innymi, modele pojazdów lub komponentów układu napędowego oraz pozwalają na opracowanie nowych modeli. Na podstawie wyników badań symulacyjnych możliwa jest analiza poziomu zużycia paliwa, energii, emisji szkodliwych związków zawartych w spalinach oraz efektywności pracy układu napędowego w profilu prędkości. W niniejszej pracy przedstawiono ocenę efektywności zastosowania napędów hybrydowych w samochodach osobowych w warunkach ruchu miasta średniej wielkości na przykładzie Kielc. Do badań symulacyjnych wykorzystano profile prędkości, zarejestrowane podczas rzeczywistych przejazdów w różnych porach dnia. Na podstawie wyników symulacji przeprowadzono analizę zużycia paliwa oraz emisji zanieczyszczeń, zarejestrowanych dla pojazd z napędem konwencjonalnym oraz pojazdów z napędem hybrydowym.

***Słowa kluczowe**: hybrydowe układy napędowe, rzeczywiste warunki jazdy, zużycie paliwa, emisja.*

1. Introduction

Vehicles equipped with hybrid drive system are becoming more and more popular. The technological solutions used in them are subject to continuous development. Both individuals and companies are increasingly willing to purchase this type of vehicles. One of the advantages of purchasing a hybrid is mainly the lower fuel consumption in comparison to conventional vehicles, which translates into lower maintenance costs [5, 15, 21].

Recognition and specification of the operating conditions of a hybrid vehicle allows for more precise estimation of fuel and energy consumption as well as emissions of harmful compounds in exhaust fumes. This is important, because the traffic intensity and type of road (city centre streets, suburban streets, highways), topography or ambient temperature affect the values of the aforementioned performance parameters. Research paper [23] features a study of the impact of the aforementioned factors on the effectiveness of using hybrid vehicles in the traffic conditions of the Quebec province (Canada). The authors collaborated with 95 vehicle owners: 74 – with conventional drive units equipped with gasoline engines and 24 – with hybrid drive units. The vehicles were fitted with instrumentation (data loggers) that recorded the instant speed, fuel consumption and idle downtimes. The data was collected for a year. The presented results demonstrated that the fuel consumption in operating conditions recorded by the hybrids were 28% lower on average than in the case of conventional vehicles. Works [2, 22] present the methodology of selection of the optimal vehicle hybrid system by studying the mobility and travel tendencies of the analysed city's residents. The data was collected by mounting GPS recorders in private vehicles for 1-18 months. The conducted studies and the analysis of the data collected allowed for estimating the average energy used per test drive. The calculations were used to determine the energy capacity of the hybrid vehicle's energy storage and distribution of the charging stations.

The ability to study the operating parameters of a vehicle in specific conditions is very troublesome. Due to the above, there is a need to use other methods of conducting the measurements. The estimation of energy consumption, emissions of harmful substances in exhaust gases or fuel consumption of the selected vehicle in specific driving conditions can be conducted during stationary tests on a chassis dynamometer or obtained as result of simulation conducted

using vehicle simulation computer programs. Firstly, it is necessary to determine the vehicle's operation conditions. One of the methods of reflecting vehicle performance is the speed profile recording during the real-world test drive. The recording is usually conducted using the GPS satellite navigation system receivers mounted in the vehicle. The obtained speed profile reflects the dynamic properties of the selected vehicle in specific conditions. The examples and methodology of conducting measurements in real-world conditions can be found in the following works [6, 12, 16].

Another method of representing the real-world traffic conditions of a specific city or region is to develop a driving cycle. It is the speed profile consisting of the sequences of acceleration, constant speed driving, braking and idling time. The cycle is substantially affected by: infrastructure, e.g. distribution and set-up of traffic lighting, type of intersections, distribution of bus stops, type of road (urban, suburban, highway), the route's vertical profile as well as the traffic intensity. The methodology of the drive cycle structures based on the speed profiles of real-world test drives is specified in a broader manner in the following works [4, 10, 11, 25]. Recorded speed profiles or developed drive cycles can be used for testing vehicles on a chassis dynamometer or in simulation testing.

The aim of this paper was to evaluate the effectiveness of hybrid passenger cars equipped with batteries of different capacities. The analysis was based on the simulation carried out using the speed profiles collected in real-world condition in various times of the day.

2. Vehicle test methods

Chassis dynamometers allow for simulating the vehicle's drive unit operating conditions in stationary conditions. The essence of the dynamometer's operation is the replacement of a static road pavement by a movable track. Smooth speed adjustment and motion resistance allows for conducting tests in steady and transient states. As result of the tests conducted with the use of the chassis dynamometer, it is possible to obtain an evaluation of the drive unit's technical condition, fuel consumption and emission of toxic compounds included in exhaust gases using additional instrumentation (e.g. the AFR – Air to Fuel Ratio sensor). The measurement can also utilize the On-Board Diagnostics. The advantage of conducting tests on a chassis dynamometer is ensuring the repeatability of measurements and the ability to check the vehicle in the conditions of particular load which is difficult to obtain during normal operation.

When conducting tests on a chassis dynamometer, it is possible to realize any selected speed profile. It allows for testing the fuel consumption and emission of harmful compounds included in exhaust gases in specific drive conditions. Using the results of tests conducted on a chassis dynamometer, the authors of papers [3,18] have compared the fuel consumption and emissions in speed profiles reflecting urban, suburban and highway driving. Research paper [9] presents fuel consumption and CO emission values recorded by a hybrid vehicle during chassis dynamometer tests. The testing was conducted based on the standard ECE-15 cycle and the Loughborough University Urban Drive Cycle (LUUDC) developed on the basis of real-world test drives. The measurement results demonstrated fuel consumption higher by 12% in the LUUDC cycle than in the ECE-15 cycle.

Vehicle testing on a chassis dynamometer allows comparison of the fuel consumption and emission of harmful substances included in exhaust gases for vehicles with various types of drive systems. For example, paper [17] presents the analysis of fuel consumption as well as CO and NO_x emissions of hybrid and conventional delivery trucks. The testing was conducted on a chassis dynamometer. The studies were carried out using the drive cycle developed on the basis of real-world operating routes of vehicles from one of the companies dealing in package deliveries in Los Angeles (USA). The presented analyses show that in the set speed profiles, hybrid vehicles demonstrate lower CO and NO_x emissions by as much as 43.9% when compared to conventional vehicles. Hybrids also demonstrate lower fuel consumption by as much as 59.8%.

During chassis dynamometer tests, it is also possible to evaluate the impact of ambient temperature of the drive system's operating parameters. Paper [8] features the estimation of the impact of temperature on energy consumption and range of an electric vehicle. The testing featured three electric passenger cars. Tests conducted on a chassis dynamometer in the temperature of -20°C demonstrated increased energy consumption of up to 9% when compared to energy consumption during testing in the temperature of +23°C. It was estimated that in the Finnish Road Cycle carried out in the temperature of -20°C, the range of an electric vehicle decreased by 51% in comparison to the range specified by the manufacturer.

Another method that allows estimating the vehicle's operating parameters in specific driving conditions is simulation testing. Computer programs and software provide the ability of modeling and simulating new solutions in vehicles drivetrain without the need to construct prototypes. A hybrid drive is a complex system that combines electrical, mechanical, electrochemical and electronic components. The tools intended for modeling and simulation of hybrid drive support complex interactions between the drive unit's mechanical and electrical elements. The functionalities of vehicle modeling and simulation programs allow for using existing solutions and available vehicle models, drive units and their elements or developing new concepts and models. The most popular programs intended for simulating hybrid vehicles are: AVL Cruise, Autonomie/PSAT, GT-SUITE, LMS AMESim, ADVISOR, DYNA4 Advanced Powertrain.

Computer simulation programs allow for analyzing the operation of particular drive unit components in the set drive profile. They allow conducting simulation testing in terms of analyzing the dynamics (e.g. ability to accelerate, overcome elevations, reach maximum speed); forecasting, evaluation and optimization of fuel consumption; analyzing the control system and diagnostics; analyzing the structure's cohesion to facilitate the testing and validation of components; preliminary evaluation and analysis of a new concept or solution; estimation of predicted emission.

The vast majority of simulation programs allows for implementing customized drive cycles. This allows analyzing the level of fuel and energy consumption as well as emissions of harmful compounds in exhaust gases and the operating effectiveness of the drive system in real-world traffic conditions. Research papers [14,19] present the comparison of fuel consumption and emission of harmful compounds included in exhaust gases. The simulation testing of vehicles with various type of drivetrain was conducted in drive cycles based on real-world test drives. Paper [7] presents the model of a hybrid commercial vehicle developed in the Autonomie program. The simulation was carried out using the speed profiles recorded during real-world operating routes. The author demonstrated the impact of the vehicle's load on the fuel consumption in the analyzed operating cycles.

Vehicle simulation tests allow for analyzing the operating effectiveness of particular drive elements. By using the vehicle simulation programs, it is possible to conduct an evaluation and determine the operating characteristics of the following:

- combustion engine,
- exhaust gases treatment system,
- cooling system,
- temperature distribution among the drive system's components,
- lubrication system,
- fuel injection system,
- hydraulic and pneumatic systems,
- analysis of the energy storage performance,
- electric engine,
- energy management system.

The paper [1] presents an analysis of the performance parameters of a combustion engine as well as the fuel consumption and CO_2 emissions of vehicles with conventional and parallel hybrid drive. The simulation was conducted for three standard cycles: UDDS, FTP and US06HWY, as well as for the drive cycle developed for the city of Baqubah (Iraq). The presented results demonstrate that the use of a hybrid drive in the driving conditions of the analyzed city can reduce fuel consumption by up to 68%. Paper [13] presents the simulation results of conventional, series hybrid and plug-in hybrid vehicle in the drive cycle developed for Kansas (USA). The authors conducted a comparative analysis of fuel consumption and the performance of the energy storage in the analyzed driving conditions. Work [24] presents an analysis of the effectiveness and the operating parameters of selected elements of a city bus' hybrid drive. The simulation testing was conducted in the AVL Cruise program by using the speed profile recorded during an real-world test drive of a city bus in Madrid (Spain). Paper [20] presents an analysis of operation of a plug-in hybrid vehicle's energy storage. The purpose of the simulation was to investigate various methods of battery charging and configuration of the energy management system.

3. Research methodology

3.1. Tests in real-world conditions

Kielce is a medium sized city located in south-central part of Poland. The measurement route went along centre streets in the city of Kielce. The length of the test route was 5.4 km. The route and its vertical profile was presented on Fig. 1. The route started in the point marked as A, went through dual carriageway streets and ended in point B. Due to the city's location in upland areas, the route was characterized by a rather substantial disparity in elevation that amounted to approx. 35 m. The route's maximum gradient of the road amounted to 6%.

The test vehicle was Ford Transit. The recording of the movement parameters was done by using measurement equipment mounted in the vehicle, consisting of:

- the S-350 Aqua Datron® optoelectronic sensor for measuring longitudinal speeds (Fig. 2a),
- the uEEP-12 Datron® data acquisition station (Fig. 2b), with the ARMS® data analysis software.
- GPS DATA LOGGER KISTLER® (Fig. 2c),
- the TAA Datron® three-directional linear acceleration sensor.

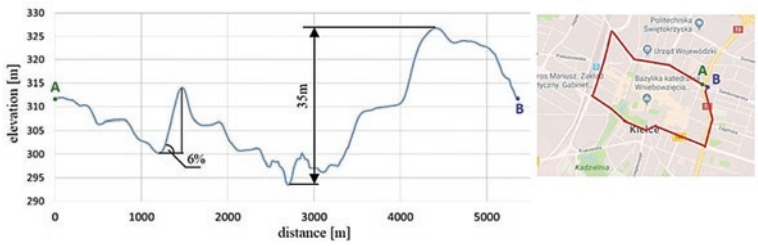


Fig. 1. Location and elevation profile of the test route

Fig. 2. Measurement equipment used to conduct the tests in real-world conditions

The test vehicle was equipped with vehicle tracking system using global positioning system (GPS) and the system for mobile communication (GSM) produced by Globtrak company. System provided detailed information of location, speed, and fuel consumption of the vehicle. Its functionalities allow management of the vehicle fleet and monitoring of the drivers.

The recording of the real-world vehicle movement parameters was conducted during test drives on a working day in four selected times of day: morning, noon, afternoon and evening. During the tests the following parameters were recorded: instantaneous speed, instantaneous acceleration and deceleration, drive time, distance travelled, instantaneous vehicle location. An exemplary speed profile, recorded during a test drive in the morning, between 7:00 - 8:00 A.M., is presented in Fig 3.

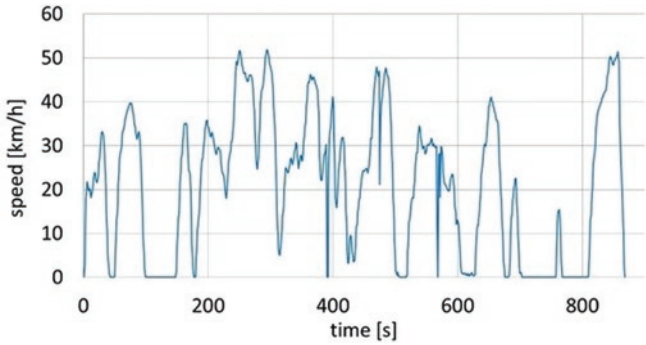


Fig. 3. Exemplary speed profile recorded during a test drive between 7:00 and 8:00 A.M.

As demonstrated on the chart in Fig. 3, in urban traffic conditions, the movement parameters (e.g. instantaneous speed) change quite substantially. The recorded speed profiles change depending on the time of day. Large traffic intensity during the morning (9:00-10:00 A.M.) and afternoon (3:00-4:00 P.M.) rush hours elongates the travelled time. Driving is more smooth during other times of day and is characterized by higher average speed. The selected parameters of the recorded test drives are presented in Tab. 1.

Based on the conducted measurement studies, it is possible to state that test drives during morning and noon hours are characterized by similar average speed and similar travelled time. During the afternoon rush hours (3:00-4:00 P.M.), the average speed is clearly lower and the share of stop phase amounts to 40% of the total time of test drive. In late afternoon (6:00-7:00 P.M.) or evening (8:00-9:00 P.M.), after the rush hours, the time of test drive is substantially shorter, which results in an increase in the average speed. The test drives are characterized by high smoothness, which is caused by lower traffic intensity. They feature an increase in average speed and the share of stop phase can constitute little more than 9% of the total time of test drive.

3.2. Simulation tests

The speed profiles recorded during real-world measurements tests were implemented into the vehicle simulation program – ADVISOR (ADvanced Vehicle SImulatOR). The program operates in the Matlab/Simulink. ADVISOR is a popular tool for simulating vehicles with various drive configurations. It was developed by the scientists from the American National Renewable Energy Laboratory (NREL). The program features built-in models of vehicles with conventional, series and parallel hybrid, electric and hydrogen cell drive.

With the use of complex database, the user develops vehicle model with the help of drop-down menus in the dialogue box. Firstly, the user selects the vehicle type, drive system and particular elements of the drive by specifying their capacity, efficiency and weight. Then, the

Table 1. Selected parameters of the recorded test drives

	7:00-8:00 A.M.	9:00-10:00 A.M.	11:00-12:00 A.M.	3:00-4:00 P.M.	6:00-7:00 P.M.	8:00-9:00 P.M.
time [s]	868	942	921	1073	773	593
average speed [km/h]	22.06	21.83	22.40	18.05	24.75	33.08
stop phase duration [s]	212	264	282	424	152	56
percentage time of stop phase in total travelled time [%]	24.40	28.00	30.60	39.50	19.70	9.40

user selects the drive cycle. With the assumed drive unit configuration and specified drive cycle, the program estimates the energy consumption and the performance of analysed type of drive train. Fig. 4 presents a parallel hybrid vehicle model developed in the ADVISOR.

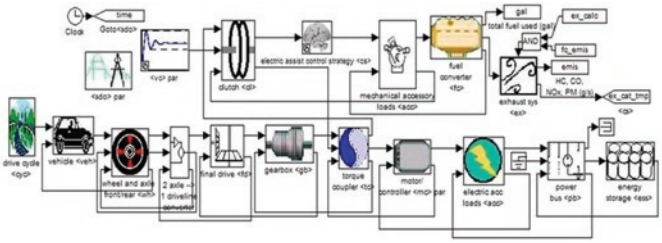


Fig. 4. Model of parallel hybrid vehicle in ADVISOR

ADVISOR allows to modify the models by importing files with the vehicle's data, characteristics and parameters of the drive components and energy storage or developing and implementing new models. It is also possible to add new drive cycle by importing files with such parameters as speed determined as a function of time or road elevation profile determined as a function of road distance.

The vehicle models available in the ADVISOR program were modified and passenger car models with conventional and parallel hybrid (HEV) drive were developed. The front area of the analysed vehicles amounts to 2.66 m^2, rolling resistance coefficient amounts to 0.009 and the aerodynamic resistance coefficient amounts to 0.44. For all simulation cases, the curb weight was 1,200 kg increased by a load of 150 kg was used. In the case of a hybrid vehicle, the weight was additionally increased by the battery weight. The selected parameters of the vehicles used in the simulation are presented in Tab. 2.

Table 2. Parameters of vehicles used in simulation tests

	Conventional	HEV			
engine power [kW]	96	74			
electric machine power [kW]	-	62			
battery capacity [kWh]	-	8,8	6,5	4,6	2,2
weigh [kg]	-	127	95	64	32

In the case of the hybrid vehicle (HEV), simulation was conducted for various capacities of energy storages. The initial battery state of charge prior to any trip amounted to 70%.

The fuel consumption results obtained from simulation of conventional vehicle were compared with data derived from vehicle monitoring system based on GPRS and GPS technology – Globtrak. The fuel consumption values acquired from ADVISOR indicated values nearly 10% higher than those given by Globtrak system.

4. Results

Based on the results of simulation of selected vehicles, the following parameters were used for further analysis: average fuel consumption and emission: PM_x, CO and NO_x. Exemplary simulation results are presented in Fig. 5 and Fig. 6. They include the instantaneous emission of PM_x, CO and NO_x as well as fuel consumption during a test drive at 3:00-4:00 P.M.

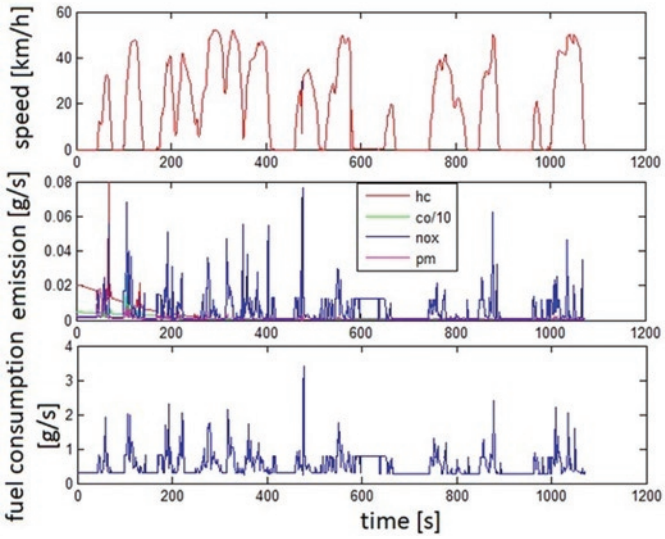


Fig. 5. Results of simulations of a conventional drive unit for the test drive at 3:00-4:00 P.M.

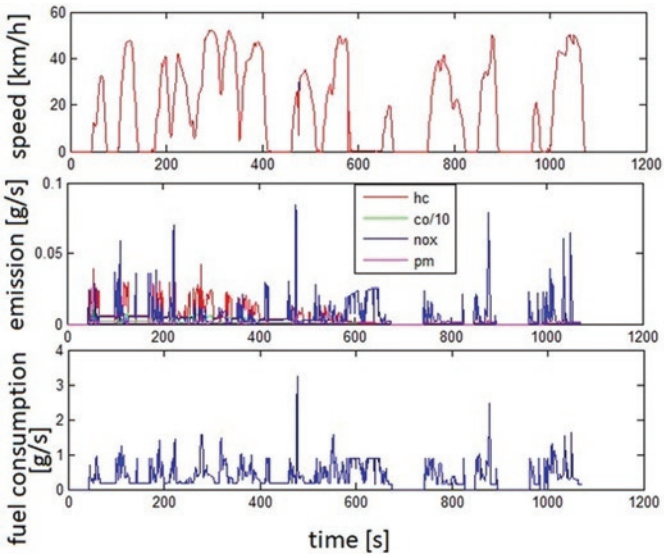


Fig. 6. Results of simulations of a hybrid drive unit with battery capacity of 8.8 kWh for the test drive at 3:00-4:00 P.M.

The average fuel consumption of the analysed vehicles is presented in Fig 7. Regardless of the time of day, the conventional vehicle recorded the highest average fuel consumption.

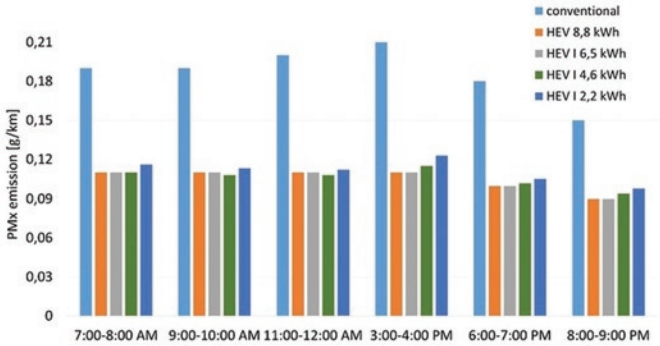


Fig. 7. Average consumption during the analysed test drive

The average fuel consumption obtained by hybrid vehicles clearly demonstrate that the higher the energy storage capacity is, the lower is the fuel consumption. In the analysed cases, the hybrid vehicle, equipped with energy storage system with the capacity of 8.8 kWh, recorded lower fuel consumption by 24% on average in relation to a conventional vehicle. It is worth noting that the differences in average fuel consumption recorded by hybrid and conventional vehicles were the highest during test drives at 3:00-4:00 P.M. It is caused by the road conditions. During the stop time, HEV using the electric engine only, did not used fuel, thereby the idling was eliminated. At that time there was no emission of harmful substances included in exhaust gases. Percentage reduction of the fuel consumption of the hybrid vehicle in comparison to a conventional vehicle in analysed test drives is presented in Tab. 3.

Fig. 8 presents the emission of particulate matter (PM_x) obtained as result of the simulations of hybrid and conventional vehicles. The highest PM_x emission during test drives in the analysed times of day were recorded for the conventional vehicle. In the case of hybrids, the values were similar in each of the analysed test drives.

The lowest emission of particulate matter was recorded during test drives in the afternoon and evening (6:00-7:00 P.M. and 8:00-9:00 P.M.). In comparison to conventional vehicles, hybrid vehicles demonstrated even 42% lower PM_x emission on average (Tab. 4). It is worth noting that the biggest differences in particulate matter emission recorded by hybrid and conventional vehicles took place during test drives at 3:00-4:00 P.M. At that time, the hybrids demonstrate lower PM_x emission by up to 48%.

Fig. 9 presents the CO emission, obtained as the simulation result, recorded for the analysed vehicles. The highest CO emission in the analysed road conditions was demonstrated by the conventional vehicle.

Table 3. Percentage reduction in hybrid vehicle fuel consumption in comparison to a conventional vehicle

	HEV 8,8 kWh	HEV 6,5 kWh	HEV 4,6 kWh	HEV 2,2 kWh
7:00-8:00 A.M.	24%	24%	22%	16%
9:00-10:00 A.M.	25%	24%	21%	16%
11:00-12:00 A.M.	24%	23%	22%	17%
3:00-4:00 P.M.	28%	26%	25%	17%
6:00-7:00 P.M.	22%	21%	20%	15%
8:00-9:00 P.M.	21%	20%	19%	15%
average	24%	23%	22%	16%

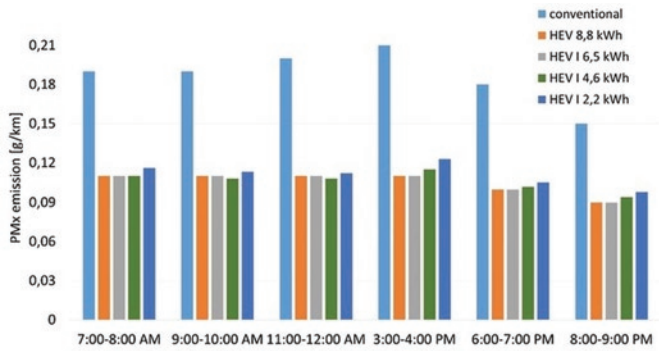


Fig. 8. PMx emission during test drives at specific times of day

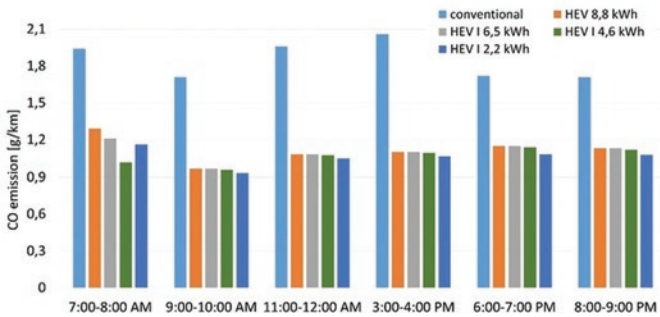


Fig. 9. CO emission during trips at specific times of day

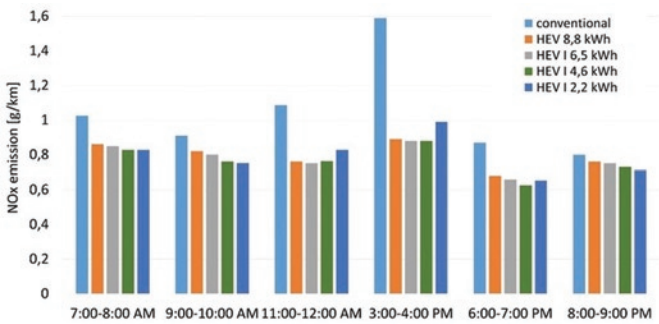


Fig. 10. NO_x emission during trips at specific times of day

In each of the analysed test drives, the hybrid vehicles demonstrated substantially lower CO emissions, regardless of the energy storage capacity. The lowest carbon oxide emissions were recorded during the afternoon test drive (Tab. 5). During the test drive at 3:00-4:00 P.M., the CO emissions recorded by the hybrids were lower by 42% in comparison to the conventional vehicle.

Fig. 10 presents the nitrogen oxides emission during test drives in the selected times of day. The conducted simulation tests demonstrate that in each of the analysed test drives, the highest NO_x emission was achieved by the conventional vehicle. The traffic conditions substantially affect the nitrogen oxide emission. This is especially clear in the case of the conventional vehicle. The NO_x emission achieved during the test drive at 3:00 – 4:00 P.M. is nearly twice as high as during the evening test drive at 8:00-9:00 P.M.

The hybrids demonstrate a 16-19% lower NO_x emission on average in comparison to the conventional vehicle (Tab. 6). It is worth noting that the nitrogen oxides emissions change depending on the traffic conditions. When comparing the NO_x emission achieved by the hybrid and conventional vehicles, the smallest differences occur during test drives with relatively small traffic intensity. The biggest differences can be observed during test drives in the afternoon rush hours (3:00-4:00 P.M.). The nitrogen

Table 4. Percentage reduction in a hybrid vehicle's PM_x emission in comparison to a conventional vehicle (conventional PMx = 100%)

	HEV 8,8 kWh	HEV 6,5 kWh	HEV 4,6 kWh	HEV 2,2 kWh
7:00-8:00 A.M.	42%	42%	42%	39%
9:00-10:00 A.M.	42%	42%	43%	41%
11:00-12:00 A.M.	45%	45%	46%	44%
3:00-4:00 P.M.	48%	48%	45%	41%
6:00-7:00 P.M.	44%	44%	43%	42%
8:00-9:00 P.M.	40%	40%	37%	35%
average	42%	42%	42%	39%

Table 5. Percentage reduction in a hybrid vehicle's CO emission in comparison to a conventional vehicle (conventional CO = 100%)

	HEV 8,8 kWh	HEV 6,5 kWh	HEV 4,6 kWh	HEV 2,2 kWh
7:00-8:00 A.M.	34%	38%	47%	40%
9:00-10:00 A.M.	43%	43%	44%	45%
11:00-12:00 A.M.	45%	45%	45%	47%
3:00-4:00 P.M.	47%	47%	47%	48%
6:00-7:00 P.M.	33%	33%	34%	37%
8:00-9:00 P.M.	34%	34%	35%	37%
Average	34%	38%	47%	40%

Table 6. Percentage reduction in a hybrid vehicle's NO_x emission in comparison to a conventional vehicle (conventional NO_x = 100%)

	HEV 8,8 kWh	HEV 6,5 kWh	HEV 4,6 kWh	HEV 2,2 kWh
7:00-8:00 A.M.	16%	17%	19%	19%
9:00-10:00 A.M.	10%	12%	16%	17%
11:00-12:00 A.M.	30%	31%	30%	24%
3:00-4:00 P.M.	44%	45%	45%	38%
6:00-7:00 P.M.	22%	24%	28%	25%
8:00-9:00 P.M.	5%	6%	9%	11%
Average	16%	17%	19%	19%

oxide emission recorded by the hybrids at that time is lower by up to 45% in comparison to the conventional vehicle.

5. Conclusion

The presented results demonstrate that the biggest differences in the emission and the average fuel consumption between the conventional and hybrid vehicles occur during the afternoon rush hours (3:00-4:00 P.M.). During that specific test drive, as much as 40% of the total test drive times are stop phases. The use of an electric engine in hybrid vehicles eliminated the idling. Thanks to this solution, hybrids do not emit harmful exhaust gases compounds during a stop phase. The presented results demonstrated that the use of a hybrid drive contributes substantially to the reduction in fuel consumption and emission. This applies especially when driving with low speed in high traffic intensity conditions.

The energy capacity of the battery used in the hybrid drive significantly affects the vehicle's performance parameters. The higher is the capacity of the energy storage devices, the bigger amount of the energy electric drive delivers for traction purposes. This translates into lower fuel consumption and emissions.

The conducted simulations confirm the possibility of verifying the effectiveness of use of a hybrid vehicle with specific parameters in relation to the specificity of a particular city traffic condition. This can facilitate specific configuration of an hybrid drive system to make its use as optimal as possible in terms of emission and fuel consumption in real-world conditions.

References

1. Al-Samari A. Study of emissions and fuel economy for parallel hybrid versus conventional vehicles on real world and standard driving cycles. Alexandria Engineering Journal 2017; 56(4): 721-726, 10.1016/j.aej.2017.04.010.
2. Björnsson L, Karlsson S. Plug-in hybrid electric vehicles: How individual movement patterns affect battery requirements, the potential to replace conventional fuels, and economic viability. Applied Energy 2015; 143: 336-347, doi:10.1016/j.apenergy.2015.01.041.
3. Fontaras G, Pistikopoulos P, Samaras Z. Experimental evaluation of hybrid vehicle fuel economy and pollutant emissions over real-world simulation driving cycles. Atmospheric Environment 2008; 42(18): 4023-4035, doi.org/10.1016/j.atmosenv.2008.01.053.

4. Galgamuwa U, Perera L, Bandara S. Developing a general methodology for driving cycle construction: comparison of various established driving cycles in the world to propose a general approach. Journal of Transportation Technologies 2015; 5: 191-203, doi:10.4236/jtts.2015.54018.
5. Hannan M, Azidin F, Mohamed A. Hybrid electric vehicles and their challenges: A review. Renewable and Sustainable Energy Reviews 2014; 29: 135-150, doi:10.1016/j.rser.2013.08.097.
6. Keramydas C, Papadopoulos G, Ntziachristos L, Lo T-S, Ng K-L,. Wong H-L A, Wong C.-L. Real-World Measurement of Hybrid Buses' Fuel Consumption and Pollutant Emissions in a Metropolitan Urban Road Network. Energies 2018; 11: 1-17, 10.3390/en11102569.
7. Lajunen A. Fuel economy analysis of conventional and hybrid heavy vehicle combinations over real-world operating routes. Transportation Research Part D 2014; 31: 70–84, 10.1016/j.trd.2014.05.023.
8. Laurikko J, Granström R, Haakana A. Realistic estimates of EV range based on extensive laboratory and field tests in Nordic climate conditions. World Electric Vehicle Journal 2013; 6: 192-203, 10.1109/EVS.2013.6914919.
9. Lintern M, Chen R, Carroll S, Walsh C. Simulation study on the measured difference in fuel consumption between real-world driving and ECE-15 of a hybrid electric vehicle. Proceedings of the Hybrid and Electric Vehicles Conference (HEVC 2013), 6-7 November 2013, London, UK, 10.1049/cp.2013.1918.
10. Lipar P, Strnad I, Česnik M, Maher M. Development of Urban Driving Cycle with GPS Data Post Processing. Promet - Traffic & Transportation 2016; 28(4): 353-364, doi.org/10.7307/ptt.v28i4.1916.
11. Mansour C, Haddad M, Zgheib E. Assessing consumption, emissions and costs of electrified vehicles under real driving conditions in a developing country with an inadequate road transport system. Transportation Research Part D: Transport and Environment 2018; 63: 498-513, doi.org/10.1016/j.trd.2018.06.012.
12. Millo F, Rolando L, Fuso R, and Zhao J. Development of a new hybrid bus for urban public transportation, Applied Energy 2015; 583-594, doi:10.1016/j.apenergy.2015.03.131.
13. Moawad A, Singh G, Hagspiel S, Fellah M, Rousseau A. Impact of real world drive cycles on PHEV fuel efficiency and cost for different power train and battery characteristics. Proceedings of the International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium (EVS24) 2009, Stavanger, Norway. 1-10, 10.3390/wevj3010186.
14. Oh Y, Park J, Leeb J, Seo J, Park S. Estimation of CO2 reduction by parallel hard-type power hybridization for gasoline and diesel vehicles. Science of The Total Environment 2018; 59: 2-12, 10.1016/j.scitotenv.2017.03.171, 2017.
15. Pawełczyk M, Szumska E. Evaluation of the efficiency of hybrid drive applications in urban transport system on the example of a medium size city. MATEC Web of Conferences 2018, 180: 1-7, https://doi.org/10.1051/matecconf/201818003004.
16. Pitanuwat S, Sripakagor A. An Investigation of Fuel Economy Potential of Hybrid Vehicles under Real-World Driving Conditions in Bangkok, Energy Procedia 2015; 79: 1046–1053, doi:10.1016/j.egypro.2015.11.607.
17. Russell R, Johnson K, Durbin T, Chen P, Tomic J, Parish R. Emissions, Fuel Economy, and Performance of a Class 8 Conventional and Hybrid Truck. SAE Int. J. Commer. Veh. 2013; 6(2): 545-554, dx.doi.org/10.4271/2013-01-2468.
18. Suarez-Bertoa R, Astorga C. Unregulated emissions from light-duty hybrid electric vehicles, Atmospheric Environment 2016: 136: 134-143, 10.1016/j.atmosenv.2016.04.021.
19. Wang H, Zhang X, Ouyang M. Energy consumption of electric vehicles based on real-world driving patterns: A case study of Beijing. Applied Energy 2015; 157: 710-719, 10.1016/j.apenergy.2015.05.057.
20. Woo D, Choe G, Kom J, Lee B, Hur J, Kang G. Comparison of integrated battery chargers for plug-in hybrid electric vehicles: Topology and control. Proceedings of the IEEE International Electric Machines & Drives Conference (IEMDC), 2011, Niagara Falls, Canada, 10,1109/IEMDC.2011.5994791.
21. Wu G, Inderbitzin A, Bening C. Total cost of ownership of electric vehicles compared to conventional vehicles: A probabilistic analysis and projection across market segments. Energy Policy 2015; 80: 196–214, doi:10.1016/j.enpol.2015.02.004.
22. Wu X, Dong J, Lin Z. Cost analysis of plug-in hybrid electric vehicles using GPS-based longitudinal travel data. Energy Policy 2014; 68: 206-217, doi:10.1016/j.enpol.2013.12.054.
23. Zahabi S, Miranda-Moreno L, Barla P, Vincent B. Fuel economy of hybrid-electric versus conventional gasoline vehicles in real-world conditions: A case study of cold cities in Quebec, Canada. Transportation Research Part D: Transport and Environment 2014; 32: 184-192, doi:10.1016/j.trd.2014.07.007.
24. Zamora R, López Martínez DJ, Loboguerrero Carrasco J, Delgado Vaca J. Development of an in-series hybrid urban bus model and its correlation with on-board testing results. World Electric Vehicle Journal 2013: 6: 405-415, 10.3390/wevj6020405.
25. Zito R, Primerano F. Drive cycle development methodology and results. Transport System Centre, Adelaide: University of South Australia, 2005.

Emilia SZUMSKA
Rafał JURECKI
Department of Automotive Engineering and Transport
Kielce University of Technology
Al. Tysiąclecia P.P. 7, 25-314 Kielce, Poland

Marek PAWEŁCZYK
Department of Production Engineering
Kielce University of Technology
Al. Tysiąclecia P.P. 7, 25-314 Kielce, Poland

E-mails: eszumska@tu.kielce.pl, rjurecki@tu.kielce.pl, m.pawelczyk@tu.kielce.pl

Article citation info:
Zhang X, Zhao J. Compound fault detection in gearbox based on time synchronous resample and adaptive variational mode decomposition. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 161–169, http://dx.doi.org/10.17531/ein.2020.1.19.

Xin ZHANG
Jianmin ZHAO

COMPOUND FAULT DETECTION IN GEARBOX BASED ON TIME SYNCHRONOUS RESAMPLE AND ADAPTIVE VARIATIONAL MODE DECOMPOSITION

WYKRYWANIE ZŁOŻONYCH BŁĘDÓW PRZEKŁADNI NA PODSTAWIE SYNCHRONICZNEGO PRÓBKOWANIA WTÓRNEGO ORAZ ADAPTACYJNEJ METODY WARIACYJNEJ DEKOMPOZYCJI MODALNEJ

Compound fault detection of gearboxes is an ambitious matter considering its interconnection and complication. An innovative means for compound fault detection based on time synchronous resample (TSR) and adaptive variational mode decomposition (AVMD) is put forward in this work. TSR used in the method can enhance fault signals of synchronous shaft gears by eliminating signal components independent of synchronous shaft. Therefore, the TSR is used to separate the synchronous shaft signal corresponding to the gear fault from the raw compound fault signal. Then a series of mode components are obtained by decomposing the synchronous shaft signals of all faults by AVMD. The variational mode decomposition (VMD) can overcome the mode aliasing problem of empirical mode decomposition (EMD), but the decomposition effect of VMD is affected by its parameter setting. Thus, the paper proposes an AVMD algorithm based on whale optimization algorithm (WOA). In the AVMD, the WOA is used to optimizes the parameters of the VMD. After AVMD decomposition, the correlated kurtosis of the mode components obtained by AVMD decomposition is calculated. Then the mode components with the maximum correlated kurtosis are selected to carry out envelope analysis. Finally, the compound fault feature can be found from the envelope spectrum to get the diagnosis results. In order to test the validity of the proposed method, a compound fault experiment is implemented in a gearbox. Through the analysis of the experimental data, it is proved that the method shows a good performance in the compound fault detection of gearbox.

Keywords: *compound fault; gearbox; time synchronous resample; adaptive variational mode decomposition.*

Wykrywanie złożonych błędów przekładni stanowi trudne zagadnienie ze względu na ich skomplikowany charakter i powiązania wewnętrzne. W pracy zaproponowano nowatorską metodę wykrywania błędów złożonych opartą na synchronicznym próbkowaniu wtórnym (TSR) oraz adaptacyjnej metodzie wariacyjnej dekompozycji modalnej (AVMD). TSR pozwala wzmacniać sygnały błędów występujących w synchronicznych przekładniach walcowych, dzięki eliminacji składowych sygnału niezwiązanych z działaniem wału synchronicznego. Dlatego też w przedstawionych badaniach, TSR wykorzystano do wyodrębnienia sygnału wału synchronicznego odpowiadającego błędowi przekładni, z surowego sygnału błędu złożonego. Następnie wszystkie sygnały błędu wału synchronicznego poddano dekompozycji za pomocą AVMD, dzięki czemu otrzymano szereg składowych modalnych. Wariacyjna dekompozycja modalna (VMD) pozwala uniknąć problemu aliasingu, który występuje w przypadku empirycznej dekompozycji modalnej (EMD), przy czym efekt dekompozycji zależy od ustawień parametrów. Dlatego w artykule zaproponowano adaptacyjny algorytm VMD oparty na algorytmie optymalizacji wielorybów (WOA), który optymalizuje parametry VMD. Następnym krokiem po dekompozycji AVMD, było obliczenie skorelowanej kurtozy składowych modalnych otrzymanych na drodze tej dekompozycji. Składniki modalne o najwyższych wartościach skorelowanej kurtozy wykorzystano do przeprowadzenia analizy obwiedni. Błąd złożony wykrywano na podstawie widma obwiedni. Skuteczność proponowanej metody sprawdzono przeprowadzając doświadczenie na przekładni, w której występował błąd złożony. Wyniki eksperymentu pokazują, że proponowane podejście stanowi skuteczną metodę wykrywania złożonych błędów.

Słowa kluczowe: *błąd złożony; przekładnia; synchroniczne próbkowanie wtórne; adaptacyjna metoda wariacyjnej dekompozycji modalnej.*

1. Introduction

Gearboxes are vital elements that are extensively used in automobile, aeroplanes and energy equipment. The gearboxes fault account for 80% in the shutdown malfunction of the transmission machinery[20, 22]. Therefore, it is essential to carry out gearbox fault diagnosis to prevent the gearbox from malfunction and reduce the economic loss[27]. Due to the long running time and poor working conditions, the faults of gearbox often occur in the form of compound fault simultaneously. Compound fault of gearbox may cause more serious consequences or unnecessary economic losses in maintenance activities. Thus, it is essential to develop the study on compound fault detection technology of gearbox. At present, the compound fault vibration signals collected from gearbox usually have the following characteristics: (a) In the original signal collected from the gearbox, the fault component belongs to the weak signal buried in the strong signal such as gear meshing component and noise; (b) Various faults may exist at the same time and interfere with each other. This causes the fault signal to be more complex and non-stationary[4]. Therefore, the above characteristics make it inconvenient to diagnose the compound fault of gearboxes.

At present, many technologies have been introduced to compound fault diagnosis of gearbox. Guo et al.[6] put forward the gear vibration model for the planetary gear compound fault detection. Nevertheless, owing to the non-linearity, non-stationarity and complexity of

the compound fault signal and the complexity of the internal structure of the gearbox, it is laborious to raise an available model for compound fault detection. Blind Source Separation (BSS) algorithm has been applied to compound fault diagnosis of rotating machinery[7, 15], but the high requirement for raw signals limits the application of BSS algorithm. The sparse decomposition[31], spectral kurtosis[1] and morphological component analysis[30, 3] have also been introduced for compound fault detection of gearbox and have shown good performance. Wavelet transform (WT) is a commonly used and effective time-frequency analysis method. So many scholars have been put forward some compound fault diagnosis methods of gearbox based on WT. Purushotham et al.[19] proposed a compound fault means for rolling bearing based on WT. Similarly, many improved wavelet transforms have been proposed and introduced for compound fault detection, such as multiwavelet transform[13, 8], multiwavelet packet transform[25, 11], dual-tree complex wavelet transform[21, 26, 28] and empirical wavelet transform[5, 12]. However, the selection of wavelet basis function will determine the performance of WT, which is also a major disadvantage of WT. As an adaptive time-frequency analysis method, empirical mode decomposition[10] (EMD) can adaptively decompose the signal into a certain number intrinsic mode functions satisfying certain conditions. EMD is suitable for processing nonlinear and non-stationary signals because of its self adaptability. For this reason, EMD is introduced into the analysis of gearbox mixed fault signal, which have shown good performance[9]. But the EMD has the disadvantage of mode aliasing, which will affect the effect of compound fault feature extraction. The ensemble empirical mode decomposition (EEMD) is an improved signal analysis method based on EMD, which can alleviate the disadvantages of modal aliasing in EMD[29]. Sandip et al.[23] have put forward a compound fault detection approach based on EEMD and Convolution Neural Networks (CNN)[33]. It is found that there is still a certain degree of modal aliasing in EEMD and the method is sensitive to the noise existing in the signal[18]. The local mean decomposition (LMD) was proposed by Jonathan S. Smith and has been used to analyze electroencephalogram signal[24]. Jiao et al. carried out multi-faults diagnosis of rotor system using LMD-based time-frequency representation. However, LMD still has a certain degree of mode aliasing, which affects the diagnostic results.

The variational mode decomposition (VMD) method is a novel adaptive signal processing method proposed by Konstantin Dragomiretskiy[32]. It can overcome some shortcomings of EMD, such as mode aliasing and endpoint effect. VMD is completely different from the recursive decomposition algorithm of EMD. Its overall framework is a constrained variational problem and has a solid theoretical foundation. The VMD suppose that each mode element is closely surrounded by a central frequency, and transforms determination of mode bandwidth into a constrained variational problem. Separation of mode elements is achieved by solving the constrained variational problem. VMD can segment the signal frequency domain flexibly and extract the latent feature information effectively. The number of mode components and penalty factors are two critical parameters of VMD, which can influence the performance of VMD in decomposition signal. The better decomposition results of VMD need appropriate parameters. Aiming at this problem, the paper proposes an adaptive variational mode decomposition (AVMD) algorithm based on whale optimization algorithm (WOA). WOA is a new intelligent optimization algorithm put forward by Mirjalili[17], which imitates the hunting strategy of humpback whales. The advantages of WOA algorithm include less parameter settings and fast optimization speed.

In addition, the raw vibration signals collected from gearbox often contain noise and other interference signals besides fault signals[14]. Compared with other interference signals, fault signals belong to weak signals. Therefore, it is necessary to preprocess the original signal and enhance the fault signal. Time synchronous average (TSA)[2] is an effective technique in signal preprocessing for gearbox. TSA can enhance fault signals of synchronous shaft gears and their meshing gears by eliminating signal components independent of synchronous shaft, such as bearing vibration, motor vibration, gear meshing vibration independent of synchronous shaft and vibration from other mechanical equipment[16]. However, TSA will filter out the bearing fault signal in gearbox, and the signal length will be greatly shortened after the average synchronization, so the algorithm has requirements on the signal length. Therefore, a signal preprocessing method based on TSR is presented in this work.

The remainder of this article is arranged as follows. The fundamental theory of the proposed method is elaborated in the Section 2. In Section 3, the procedure of AVMD is presented. The flow chart of compound fault detection approach based on TSR-AVMD is illustrated in the Section 4. In Section 5, the performance of TSR-AVMD is validated by using compound fault experimental data of gearbox, and results are compared with other methods. The conclusions are given in the Section 6.

2. Materials and Methods

2.1. Time synchronous resample

For continuous signal $g(t)$, if it satisfies the Dirichlet condition, its Fourier transform is shown as follows:

$$X(\omega) = \int_{-\infty}^{+\infty} g(t)\exp(-i\omega t)dt \tag{1}$$

Further assume that the signal is a limited bandwidth signal:

$$X(\omega) = 0 \quad |\omega| > 2\pi f_g \tag{2}$$

The sampling frequency band $\{z(p)\}$ can be obtained by sampling the signal $g(t)$ at the starting time t_z with the sampling frequency $f_s > 2f_g$:

$$z(p) = g(\frac{P}{f_s} - t_z) \qquad p = 0,1,\cdots,p-1 \tag{3}$$

where t_z is the beginning time of sampling.

N data segments of R sampling points can be obtained through the same method:

$$x_n(r) = g(\frac{r}{f_s} - t_n) \qquad r = 0,1,\cdots,R-1 \qquad n = 0,1,\cdots,N-1 \tag{4}$$

where t_n is the sampling beginning time of the nth sampling data segment, and N is called the average segment number. If it is assumed that the start times of these data segments correspond to the same signal flag, and these data segments are synchronized, then new data segments can be obtained on average for these data segments:

$$y(r) = \frac{1}{N}\sum_{n=0}^{N-1} x_n(r) \qquad r = 0,1,\cdots,R-1 \tag{5}$$

where $y(r)$ is called time synchronized averaging signal, and the above procedure is called time synchronous average. But after TSA processing, the signal related to bearing fault will be filtered out, and the bearing fault in gearbox cannot be detected. In addition, the length of signal is greatly reduced after TSA processing, which affects the

subsequent analysis. For this reason, the input signal of TSA needs enough length, which limits the application of this technology. Therefore, this paper only carries out synchronous resampling without average processing to overcome these shortcomings, which is called time synchronous resample.

2.2. Variational mode decomposition

The theory of VMD will be illustrated in this section. The intrinsic mode function in VMD refers to an amplitude modulation-frequency modulation signal, which is shown as follows:

$$u_k(t) = A_k(t)\cos(\varphi_k(t)) \tag{6}$$

where $A_k(t)$ indicates the signal amplitude, $\varphi_k(t)$ is the phase of the signal. $\omega_k(t)=\frac{d\varphi(t)}{dt}$ stands for instantaneous frequency. The mode mentioned here is assumed to be a finite bandwidth component with a central frequency. The VMD is to seek the intrinsic mode function, which the sum of K estimation bandwidth is the smallest. The constraint condition is the sum of IMFs equal to the primary signal $f(t)$. The specific measures to constructing constrained variational models are shown below:

Hilbert transform is performed for each IMF, as shown in the following formula:

$$\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t) \tag{7}$$

For each IMF component, the corresponding center frequency ω_k is estimated and multiplied with the exponent signal $e^{-j\omega_k t}$, and the corresponding mode spectrum of each fundamental band is modulated:

$$((\delta(t)+\frac{j}{\pi t})*u_k(t))e^{-j\omega_k t} \tag{8}$$

The square of the norm of the gradient L^2 above the modulation signal is calculated. The bandwidth of each IMF component is evaluated. The following constraint variational model is constructed:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \{\sum\limits_{k=1}^{K}\left\|\partial_t\left[(\delta(t)+\frac{j}{\pi t})*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\} \\ s.t.\ \sum\limits_{k=1}^{K}u_k(t)=f(t) \end{cases} \tag{9}$$

where $\delta(t)$ is unit impulse function, j is imaginary unit, * indicates the convolution operation, ∂_t indicates partial derivation of functions, $\{u_k\}=\{u_1,u_2,\cdots,u_K\}$ indicates the decomposed K IMFs elements. $\{\omega_k\}=\{\omega_1,\omega_2,\cdots,\omega_K\}$ represents the central frequency of each IMF component.

To solve the problem of Eq (9), penalty factor α and Lagrange multiplier λ are introduced to transform the constrained variational problems into unconstrained ones. The augmented Lagrange expression is obtained in the following form:

$$\begin{aligned} L(\{u_k\},\{\omega_k\},\lambda) := \alpha\sum_{k=1}^{K}\left\|\partial_t\left[(\delta(t)+\frac{j}{\pi t})*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2 \\ +\left\|f(t)-\sum_{k=1}^{K}u_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_{k=1}^{K}u_k(t)\right\rangle \end{aligned} \tag{10}$$

The alternate direction method of multipliers are used to iteratively update u_k, ω_k and λ to search the saddle point of augmented Lagrange expression. Specific implementation steps are shown as follows:

(1) Initialize the $\{\hat{u}_k^1\},\{\omega_k^1\},\hat{\lambda}^1,n$
(2) Repeat cycle: n=n+1
(3) For all $\omega\geq 0$, update the $\hat{u}_k,\omega_k,\hat{\lambda}$

$$\hat{u}_k^{n+1}(\omega)=\frac{\hat{f}(\omega)-\sum\limits_{i=1}^{k-1}\hat{u}_i^{n+1}(\omega)-\sum\limits_{i=k+1}^{K}\hat{u}_i^{n}(\omega)+\frac{\hat{\lambda}^n(\omega)}{2}}{1+2\alpha(\omega-\omega_k^n)^2} \tag{11}$$

$$\omega_k^{n+1}=\frac{\int_0^\infty\omega\left|\hat{u}_k^{n+1}(\omega)\right|^2d\omega}{\int_0^\infty\left|\hat{u}_k^{n+1}(\omega)\right|^2d\omega} \tag{12}$$

$$\hat{\lambda}^{n+1}(\omega)=\hat{\lambda}^n(\omega)+\tau(\hat{f}(\omega)-\sum_{k=1}^{K}\hat{u}_k^{n+1}(\omega)) \tag{13}$$

(4) Repeat the (2) and (3) steps unless the iteration termination condition is met.

$$\sum_{k=1}^{K}(\left\|\hat{u}_k^{n+1}(\omega)-\hat{u}_k^n(\omega)\right\|_2^2\Big/\left\|\hat{u}_k^n(\omega)\right\|_2^2)<\varepsilon \tag{14}$$

End the iteration and get K IMFs components.

2.3. Theory of whale optimization algorithm

The whale optimization algorithm is an innovative intelligent optimization algorithm which is mainly formed by simulating the process of humpback whale's predation. The predatory behavior of humpback whales can be summarized as the following three behaviors: randomly searching for prey, surrounding target prey and preying on target prey. In WOA, the position of each humpback whale is expressed as a feasible solution to the research problem.

2.3.1. Randomly searching for prey

Searching for a feasible solution to a problem can be modeled on the process of whale swarm randomly searching for target prey. The mathematical model is as follows:

$$\mathbf{X}_{j+1}=\mathbf{X}_{rand}-\mathbf{A}\times\mathbf{D} \tag{15}$$

$$\mathbf{D}=\left|\mathbf{C}\times\mathbf{X}_{rand}-\mathbf{X}_j\right| \tag{16}$$

where j is the current number of iterations, **A** and **C** are coefficient vectors, $\mathbf{X}_{rand}$ is the position vector randomly selected from the current whale group, which is the possible solution.

The **A** and **C** in Eq. (15) and Eq. (16) can be obtained as follows:

$$\mathbf{A}=2\mathbf{a}\times\mathbf{r}_1-\mathbf{a} \tag{17}$$

$$\mathbf{C} = 2\mathbf{r}_2 \tag{18}$$

where **a** is a vector that falls linearly from 2 to 0, the $\mathbf{r}_1$ and $\mathbf{r}_2$ are random vectors in the range 0 to 1.

2.3.2. Surrounding target prey

The process of humpback whales approaching the target prey can be seen as the process of approaching the feasible solution in the algorithm. If the target prey is the best individual location for the current population, the location will be updated as follows:

$$\mathbf{X}_{j+1} = \mathbf{X}_j - \mathbf{A} \times \mathbf{D} \tag{19}$$

$$\mathbf{D} = \left| \mathbf{C} \times \mathbf{X}_j^* - \mathbf{X}_j \right| \tag{20}$$

where $\mathbf{X}_j$ is the position vector of current whale, $\mathbf{X}_j^*$ is currently the best whale position vector.

2.3.3. Preying on target prey

Humpback whales prey on the target through the following two strategies:

1. Shrinking encircling mechanism: This mechanism is realized by reducing the value of a, where a is a random value between [- 2, 2]; When a is in the range of [- 1,1], the position the whales are looking for is the position of the target prey. At this time, the whale group is close to the target prey, on the contrary, whales stay away from the prey.
2. Spiral updating position: The humpback whales approach their prey in a spiral motion. According to the motion mode, a mathematical model can be constructed as follows:

$$\mathbf{X}_{j+1} = \mathbf{D}' e^{bl} \cos(2\pi l) + \mathbf{X}_j^* \tag{21}$$

where $\mathbf{D}' = \left| \mathbf{X}_j^* - \mathbf{X}_j \right|$ is the distance between the current best position of the *i*th whale group and its prey, *b* is a constant for defining the shape of the logarithmic spiral, *l* is a random number in [-1, 1].

The above two mechanisms are carried out at the same time in the process of whale predation. In order to simulate this situation, a 50% probability is selected between them to update the position of the whale group. It can be realized by the following mathematical model:

$$\mathbf{X}_{j+1} = \begin{cases} \mathbf{X}_j - \mathbf{A} \times \mathbf{D} & p < 0.5 \\ \mathbf{D}' e^{bl} \cos(2\pi l) + \mathbf{X}_j^* & p \geq 0.5 \end{cases} \tag{22}$$

where *p* is a random number in [0, 1].

2.4. Adaptive variational mode decomposition

The procedure of AVMD is illustrated in this section. According to the description of VMD in the previous section, the number of components *K* and the penalty factor *α* are two critical parameters that influence the decomposition effect of VMD. If the value of K is much smaller than the number of natural modes of the signal, all modes in the signal will not be separated completely. On the contrary, some modal components in the signal may be over decomposed and finally some non-existent modes will appear. The penalty factor *α* mainly affects the bandwidth of the mode components decomposed by VMD. If the penalty factor is too small, the spectrum of component will be very wide, and the mode aliasing problem will occur easily. Conversely, the bandwidth of the component is narrowed, and the information contained in the mode component may be insufficient. At present, the determination of the above two parameters mainly depends on human subjective experience, which may lead to unsatisfactory decomposition effect of VMD. Therefore, this paper proposes an AVMD, which employs the WOA to optimize the parameters of VMD.

A fitness function must be determined when using WOA to optimize influence parameters of VMD. The fitness function values under different parameters are calculated, and the influence parameters are selected and updated by comparing fitness function values. The fitness function is the maximum correlation kurtosis of modes received by VMD decomposition. Correlated kurtosis can detect the existence of periodic impact signals. In engineering practice, the original signal collected from the equipment contains some noise impact signals, which is not periodic. However, the traditional kurtosis can only reflect the impact characteristics of the signal. The traditional kurtosis may reflect only the impact signal of the noise rather than the fault impact signal. Therefore, the correlated kurtosis is used to select the IMF which contains fault impact component. The first order correlated kurtosis can be computed as follows:

$$CK_1(\tau) = \frac{\sum_{t=1}^{N} (y(t)y(t-\tau))^2}{(\sum_{t=1}^{N} y(t)^2)^2} \tag{23}$$

where the $y(t)$ is the vibration signal, τ represents the sampling point length corresponding to the fault frequency to be detected. The *M* orders correlated kurtosis can be get as follows:

$$CK_M(\tau) = \frac{\sum_{t=1}^{N} (\prod_{m=0}^{M} y(t-m\tau))^2}{(\sum_{t=1}^{N} y(t)^2)^{M+1}} \tag{24}$$

The first order correlation kurtosis is mainly suitable for detecting early faults, and high order correlation kurtosis is mainly suitable for detecting serious faults.

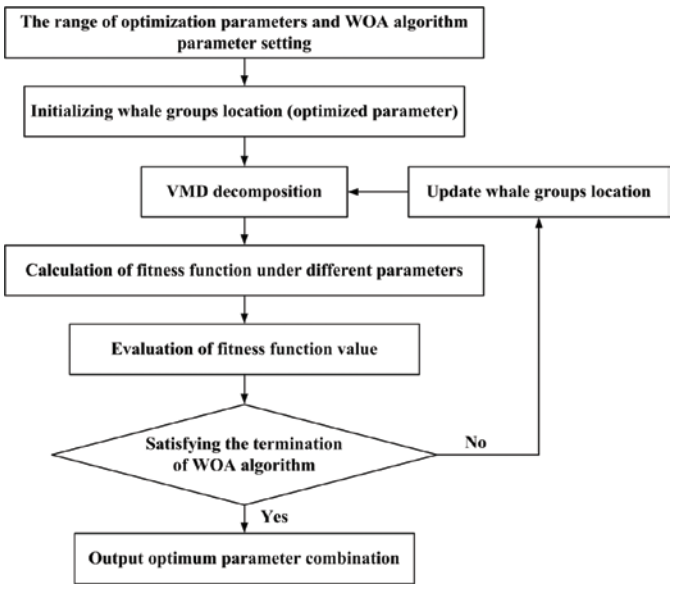


Fig. 1. Frame diagram of the AVMD

The Fig. 1 is presented the flow path of AVMD. Firstly, the optimization range for parameters and WOA algorithm parameters are set, and then the whale groups locations (optimization parameters) are initialized. Then VMD decomposition is performed to get a series of mode components. Then the fitness function is calculated, and the optimal results are chosen by evaluating fitness function. Finally, output optimization parameters when the termination condition of WOA is satisfied, otherwise the parameters are updated to continue the above operation.

2.5. Compound fault diagnosis method based on TSR-AVMD

A compound fault detection approach based on TSR-AVMD is presented in this work. The process of the proposed approach is illustrated in this section. The compound fault signals collected from gearbox include not only compound fault components, but also motor vibration signal, gear meshing vibration signal and vibration signal from other machinery and equipment. These irrelevant signals will affect the efficiency of compound fault diagnosis, so the TSR is introduced to process the original signal in this paper. TSR can enhance the fault signal of synchronous shaft gear and its meshing by removing the frequency component independent of synchronous shaft, which makes it easy to detect the fault located in synchronous shaft gear. To get the fault features and overcome the limitations of traditional VMD methods, AVMD is put forward to extract gear fault characteristic. The flow chart of proposed method for hybrid fault detection is shown as Fig. 2. The specific steps are described as follows.

(1) Collect the original signal from the gearbox through the vibration acceleration sensor.
(2) The original vibration signal is preprocessed by TSR technology, which enhances the synchronous shaft signal of each fault and eliminates the interference of the non-synchronous shaft signal component.
(3) A series of mode components are obtained by decomposing the synchronous shaft signals of all faults by AVMD.
(4) Calculate the correlated kurtosis of the mode components obtained by AVMD decomposition for all the fault synchronous shaft signals. Then the mode components with the maximum correlated kurtosis are selected for the next step.
(5) Finally, the envelope analysis of mode components with maximum correlation kurtosis is performed, and the envelope spectrum is obtained to realize fault detection.

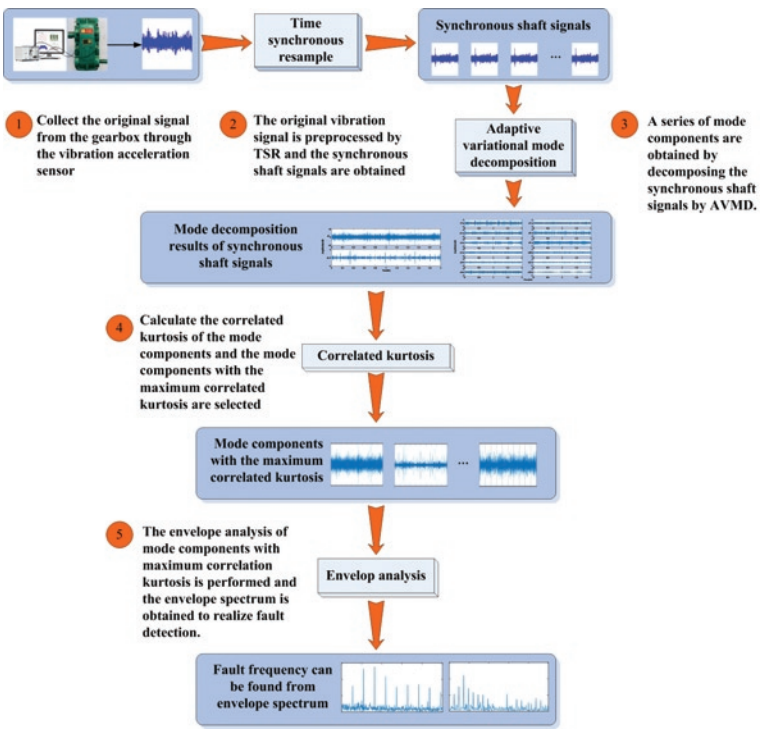


Fig. 2. Procedure of proposed method for compound fault detection

3. Experimental analysis

The performance of the proposed method is testified by using experiment signal of compound fault of gearbox in this section. The setup of the experiment is illustrated in the follows.

3.1. Experimental setup

The structure of the experimental platform is shown in the Fig. 3. The experimental data is collected by the acceleration sensor installed on the gearbox. The structure of the test gearbox and the specific layout of four acceleration sensors are shown as Fig 4. In this experiment, there are 1mm crack fault in Gear 1 and 2mm broken tooth fault in Gear 2 which is preset in the gearbox. The location of the gear fault is shown in the Fig. 5. In the experiment, the motor speed is set to 1200rpm and the load is set to 20nm. The sampling frequency of data is 20kHz.

3.2. Experimental result analysis

The experimental data of 1mm crack fault in Gear 1 and 2mm broken tooth fault in Gear 2 is employed to test the validity of proposed method. The wave form of raw compound fault signal is presented in the Fig 6. The rotational speed of input shaft is 2000rpm. According to the gear parameters shown in the Fig.4, the rotational frequency of output shaft and intermediate shaft is 2.43Hz and 10.94Hz respectively. The two gears with crack fault and broken tooth fault are located in the output shaft and intermediate shaft respectively, so the corresponding fault frequencies of the two faults are 2.43Hz and 10.94Hz respectively.

The proposed approach is employed to perform the compound fault data. Firstly, the original data is preprocessed by TSR. The interference of signal components independent of synchronous axis is eliminated, and the synchronous shaft signals corresponding to crack fault and broken tooth fault are obtained respectively. Then the WOA is employed find the optimal parameters of VMD. The parameter settings of WOA are shown in the Table 1. Fig 7 shows the change curve of fitness function during the iteration of parameter optimization with WOA. As can be seen from the Fig 7 (a), the fitness is stable when the number of iterations reaches 7 generations, which shows that the optimal solution is found for broken tooth fault signal. Similarly, the optimal solution for crack fault is found when the number of iterations reaches 16 generations as shown in the Fig 7 (b). Thus, the optimal parameters of VMD are found, and the result is presented in Table 2. VMD decomposition of synchronous shaft signal corresponding to crack fault and broken tooth fault is carried out respectively. The VMD decomposition result of synchronous shaft signal corresponding to broken tooth fault and crack fault is presented in the Fig 8 and Fig 9 respectively. Then the correlated kurtosis of all mode components is computed and the modes with maximum correlated kurtosis are chosen for envelope analysis. The IMF with maximum 15 orders correlated kurtosis for synchronous shaft signal corresponding broken tooth fault is the second IMF, and the IMF with maximum 4 orders correlated kurtosis for synchronous shaft signal corresponding gear crack fault is the first IMF. The envelope spectrums of above two IMFs are obtained as presented in Fig 10.

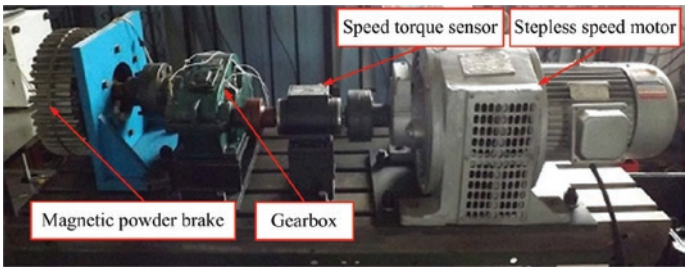


Fig. 3. Structure diagram of test rig

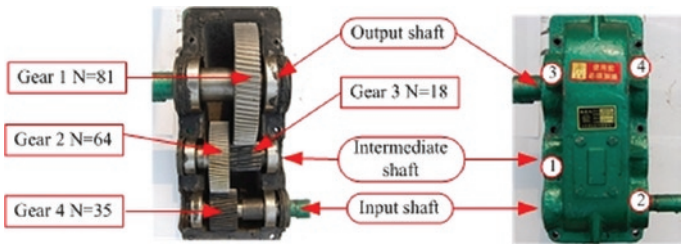


Fig. 4. Gearbox structure and sensor position

Fig. 5. Location of compound fault in gears

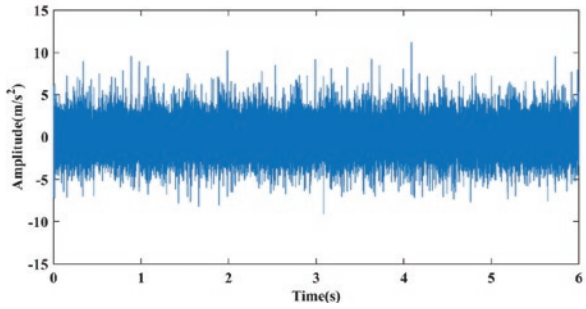


Fig. 6. The wave form of raw signal

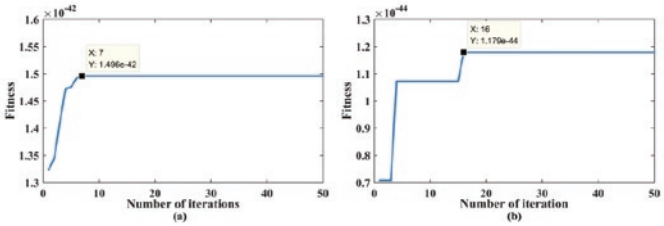


Fig. 7. The iterative of process of optimization (a) iterative process for broken tooth fault signal (b) iterative process for crack fault signal

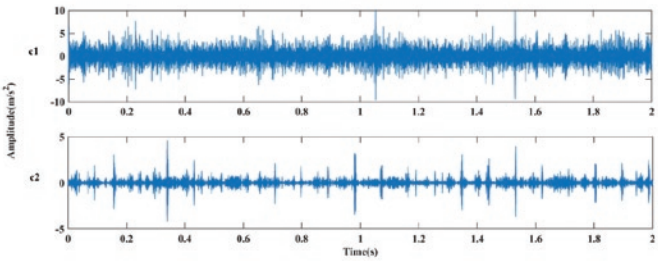


Fig. 8. Decomposition result of AVMD for synchronous shaft signal corresponding broken tooth fault

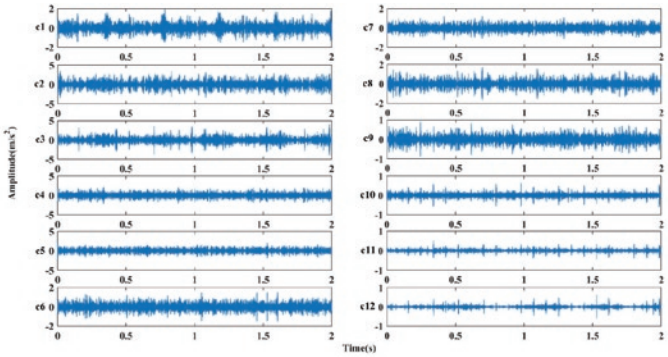


Fig. 9. Decomposition result of AVMD for synchronous shaft signal corresponding gear crack fault

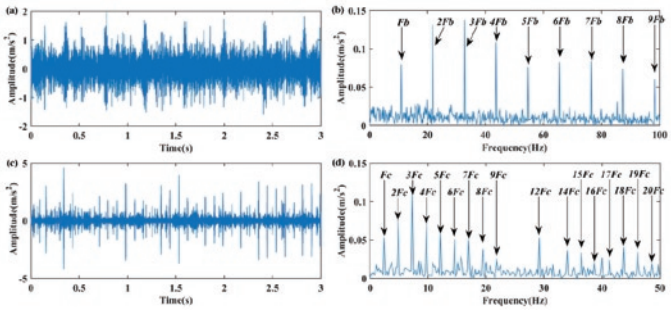


Fig. 10. The diagnosis results obtained by proposed method (a) IMF of signal with broken tooth fault (c) IMF of signal with gear crack fault; (b) & (d) The envelop spectrum of (a) and (c)

As shown in the Fig 10 (b), the prominent peak related to the fault

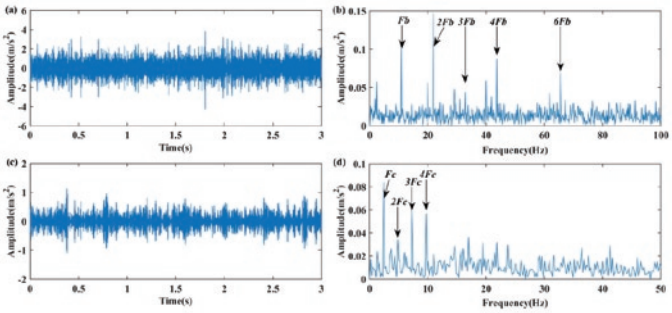


Fig. 11. Diagnosis results obtained by TSR-EMD (a) IMF of signal with broken tooth fault (c) IMF of signal with gear crack fault; (b) & (d) Envelop spectrum of (a) and (c)

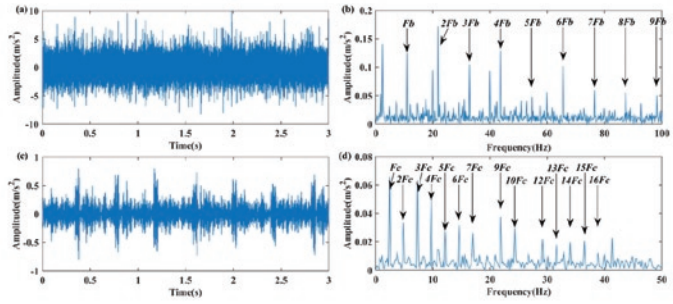


Fig. 12. Diagnosis results obtained by TSR-EEMD (a) IMF of signal with broken tooth fault (c) IMF of signal with gear crack fault; (b) & (d) Envelop spectrum of (a) and (c)

Table 1. Initial parameter settings of WOA

Number of search agents	Maximum generations	Number of parameters to be optimized	Floor of parameter K	Toplimit of parameter K	Floor of parameter α	Toplimit of parameter α
100	50	2	2	15	1000	10000

Table 2. Optimal parameters of VMD for different fault

K of VMD for broken tooth fault	K of VMD for gear crack fault	α of VMD for broken tooth fault	α of VMD for gear crack fault
2	12	7248	9247

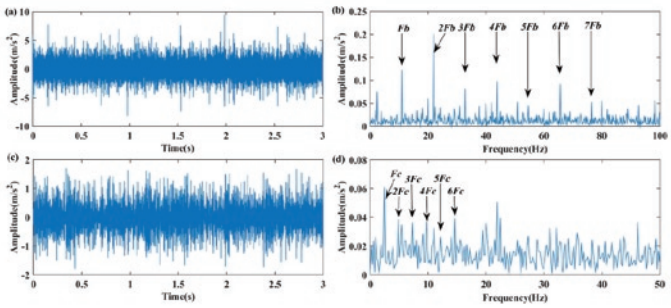


Fig. 13. Diagnosis results obtained by TSR-LMD (a) IMF of signal with broken tooth fault (c) IMF of signal with gear crack fault; (b) & (d) Envelop spectrum of (a) and (c)

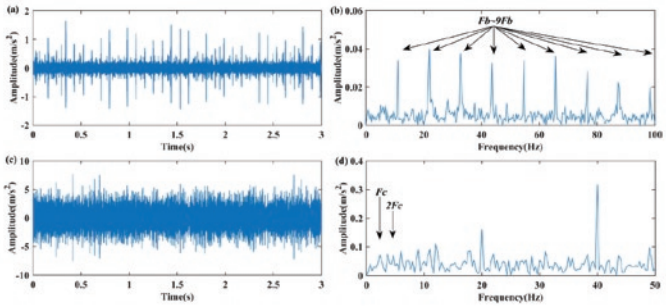


Fig. 14. Diagnosis results obtained only using AVMD (a) IMF of signal with broken tooth fault (c) IMF of signal with gear crack fault; (b) & (d) Envelop spectrum of (a) and (c)

frequency of broken tooth (*Fb*) and its harmonics can be identified evidently. As shown in the Fig 10 (d), the prominent peak corresponding to the fault characteristic frequency of crack (*Fc*) and its harmonics are evident clearly. In conclusion, the TSR-AVMD is able to separate the compound fault of broken tooth and gear crack in different shaft, and the corresponding fault can be detected effectively.

4. Discussion

In order to prove that the performance of TSR-AVMD is better than that of traditional methods, the EMD, EEMD and LMD are employed to perform the same data. Firstly, the compound fault signal is decomposed into a series of IMFs by using EMD, EEMD and LMD respectively. Then the IMFs with maximum 15 orders correlated kurtosis and maximum 4 orders correlated kurtosis are selected from the decomposition results. Finally, the selected IMFs which contains different gear fault are carried out envelope analysis and the envelope spectrums are shown in the Fig 11, Fig 12 and Fig 13.

As for envelop spectrum shown in the Fig 11 (b), the peak related to *Fb* is identified clearly. Partial harmonics of *Fb* appear in envelope spectrum, but compared to Fig 10 (b), the harmonic of some *Fb* is annihilated by other interference components. As for envelop spectrum presented in the Fig 11 (d), the prominent peak related to the *Fc* and its harmonics are identified, but not clearly. In the envelop spectrum, there are many interference components around *Fc* and its harmonics compared with Fig 10 (d). It shows that the performance of TSR-AVMD is more competitive than the TSR-EMD.

As shown in the Fig 12 (b), the peak related to *Fb* and its harmonics are identified. But in the spectrum, there are some interference components around *Fb* and its harmonics compared with the Fig 10 (b). In the Fig 12 (d), the peak related to *Fc* and its harmonics are identified clearly. It shows that the TSR-EEMD can detect the gear crack fault clearly. Nevertheless, the fault frequency of broken tooth cannot be extracted clearly enough, there are some interference components around *Fb* and its harmonics.

As for envelop spectrum shown in the Fig 13 (b), the peak related to *Fb* and its harmonics can be found, but there are disturbing components near the harmonics. It can be seen from the Fig 13 (d), the peak related to *Fc* and its harmonics are disturbed seriously. It shows that the TSR-LMD can detect the broken tooth fault, and the performance of TSR-AVMD is more competitive than the TSR-LMD. But the TSR-LMD cannot detect the gear crack fault effectively.

In conclusion, the performance of TSR-AVMD in compound fault detection is more competitive than the TSR-EMD, TSR-EEMD and TSR-LMD.

To prove the necessity of TSR in the presented approach, the AVMD is employed alone to analyze the raw signals that are not processed by TSR. Then the envelop spectrums of IMFs with 15 orders maximum correlated kurtosis and 4 orders maximum correlated kurtosis are shown as Fig 14. As shown in the Fig 14 (b), the prominent peak related to the fault frequency of broken tooth (*Fb*) and its harmonics are clearly identified. However, the prominent peak related to the characteristic frequency of crack (*Fc*) and its harmonics cannot be found in the Fig 14 (d) clearly. Thus, the gear broken tooth fault can be detected and the gear crack fault cannot be detected. In summary, it shows that the gear crack fault cannot be detected by only using AVMD, and it proves the necessity of the TSR in the proposed method.

5. Conclusions

An innovative compound fault detection approach based on TSR and AVMD is presented in this paper. In the implementation of the presented approach, the TSR is used to preprocess the raw signal to eliminate the interference of asynchronous shaft signal. Then the AVMD is employed to process the fault synchronous shaft signals obtained by TSR to extract fault features. The AVMD introduces WOA to optimize the main parameters of VMD, which overcomes the problem that the decomposition effect of VMD is affected by parameters. Then the optimal mode components that represent the fault features of gears are selected based on the principle of the maximum correlated kurtosis. Finally, the compound fault features can be extracted from the envelop spectrum of the optimal mode components. The compound fault experiment of gearbox is performed to test the validity of the TSR-AVMD. After the analysis and comparison of the experimental results, the following conclusions can be obtained.

(1) TSR is an available approach for extracting synchronous shaft fault signals and eliminating other interference signals. The experimental results show that TSR can eliminate the interference of non-synchronous shaft signal and enhance the fault signal of gearbox.
(2) AVMD can effectively overcome the shortcomings of mode aliasing in EMD. Through comparative analysis of experimental results, it can be proved that the performance of extracting fault features by AVMD is more competitive than traditional time frequency analysis methods such as EMD, EEMD and LMD.
(3) Through the experiment of compound faults in gearbox, the compound fault detection approach based on TSR and AVMD presented in this work can detect compound faults of gearbox effectively.

References

1. Antoni Jérôme, Randall R B. The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines, Mechanical Systems and Signal Processing 2006; 20(2): 308-331, https://doi.org/10.1016/j.ymssp.2004.09.002.
2. Bechhoefer E, Kingsley M. A Review of time synchronous average algorithms, Annual Conference of the Prognostics and Health Management Society 2009; 1-10.
3. Bobin Jérôme, Starck Jean-Luc, Fadili Jalal M, Moudden Yassir and Donoho David L. Morphological Component Analysis: An Adaptive Thresholding Strategy, IEEE Transactions on Image Processing 2007; 16(11): 2675-2681, https://doi.org/10.1109/TIP.2007.907073.
4. Garcia R A G, Osornio R R A and Granados L D. Smart sensor for online detection of multiple-combined faults in VSD-Fed induction motors, Sensors 2012; 12: 11989-12005, https://doi.org/10.3390/s120911989.
5. Gilles Jérôme. Empirical wavelet transform, IEEE Transactions on Signal Processing 2013; 61(16): 3999-4010, https://doi.org/10.1109/TSP.2013.2265222.
6. Guo Y C, Parker R G. Purely rotational model and vibration modes of compound planetary gears, Mechanism & Machine Theory 2010; 45: 365-377, https://doi.org/10.1016/j.mechmachtheory.2009.09.001.
7. Haile Mulugeta A, Dykas Brian. Blind source separation for vibration-based diagnostics of rotorcraft bearings, Journal of Vibration and Control 2015; 22(18): 3807-3820, https://doi.org/10.1177/1077546314566041.
8. He S, Chen J, Zhou Z et. al. Multifractal entropy based adaptive multiwavelet construction and its application for mechanical compound-fault diagnosis, Mechanical Systems and Signal Processing 2016; 76-77: 742-758, https://doi.org/10.1016/j.ymssp.2016.02.061.
9. Henriquez Patricia, Alonso Jesus B, Ferrer Miguel A, Travieso Carlos M. Review of Automatic Fault Diagnosis Systems Using Audio and Vibration Signals, IEEE Transactions on Systems, Man, and Cybernetics: Systems 2013; 44(5): 642-652, https://doi.org/10.1109/TSMCC.2013.2257752.
10. Huang N E, Shen Z, Long S R et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings Mathematical Physical & Engineering Sciences 1998; 454: 903-995, https://doi.org/10.1098/rspa.1998.0193.
11. Jiang H, Li C and Li H. An improved EEMD with multiwavelet packet for rotating machinery multi-fault diagnosis, Mechanical Systems and Signal Processing 2013; 36(2): 225-239, https://doi.org/10.1016/j.ymssp.2012.12.010.
12. Jiang Y, Zhu H, Li Z. A new compound faults detection method for rolling bearings based on empirical wavelet transform and chaotic oscillator, Chaos, Solitons & Fractals 2016; 89: 8-19, https://doi.org/10.1016/j.chaos.2015.09.007.
13. Khadem S E, Rezaee M. Development of vibration signature analysis using multiwavelet systems, Journal of Sound and Vibration 2003; 261(4): 613-633, https://doi.org/10.1016/S0022-460X(02)00992-6.
14. Li X, Li J, He D, Qu Y. Gear pitting fault diagnosis using raw acoustic emission signal based on deep learning, Eksploatacja i Niezawodnosc-Maintenance and Reliability 2019; 21(3): 403-410, https://doi.org/10.17531/ein.2019.3.6.
15. Li Z, Yan X, and Tian Z et al. Blind vibration component separation and nonlinear feature extraction applied to nonstationary vibration signals for the gearbox multi-fault diagnosis, Measurement 2013; 46: 259-271, https://doi.org/10.1016/j.measurement.2012.06.013.
16. McFadden P D. Interpolation techniques for time domain averaging of gear vibration, Mechanical Systems and Signal Processing 1989; 3: 87-97, https://doi.org/10.1016/0888-3270(89)90024-1.
17. Mirjalili S, Lewis A. The whale optimization algorithm, Advances in Engineering Software 2016; 95: 51-67, https://doi.org/10.1016/j.advengsoft.2016.01.008.
18. Pan H Y, Yang Y, Li X et al. Symplectic geometry mode decomposition and its application to rotating machinery compound fault diagnosis. Mechanical Systems and Signal Processing 2019; 114: 189-211, https://doi.org/10.1016/j.ymssp.2018.05.019.
19. Purushotham V, Narayanan S, Prasad Suryanarayana A N. Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition, NDT & E International 2005; 38 (8): 654-664, https://doi.org/10.1016/j.ndteint.2005.04.003.
20. Rashid H S J, Place C S, Mba D, Keong R L C, Healey A, Kleine-Beek W, Romano M. Reliability model for helicopter main gearbox lubrication system using influence diagrams, Reliability Engineering & System Safety 2015; 159: 50-57, https://doi.org/10.1016/j.ress.2015.01.021.
21. Seshadrinath Jeevanand, Singh Bhim, Panigrahi Bijaya Ketan. Investigation of Vibration Signatures for Multiple Fault Diagnosis in Variable Frequency Drives Using Complex Wavelets, IEEE Transactions on Power Electronics 2014; 29(2): 936-945, https://doi.org/10.1109/TPEL.2013.2257869.
22. Singh Amandeep, Parey Anand. Gearbox fault diagnosis under non-stationary conditions with independent angular re-sampling technique applied to vibration and sound emission signals, Applied Acoustics 2019; 144(15): 11-22, https://doi.org/10.1016/j.apacoust.2017.04.015.
23. Singh Sandip Kumar, Kumar Sandeep, Dwivedi J P. Compound fault prediction of rolling bearing using multimedia data, Multimedia Tools and Applications 2017; 76(18): 18771-18788, https://doi.org/10.1007/s11042-017-4419-1.
24. Smith Jonathan S. The local mean decomposition and its application to EEG perception data, Journal of the Royal Society Interface 2005; 2: 443-454, https://doi.org/10.1098/rsif.2005.0058.
25. Tabrizi A, Garibaldi L, Fasana A, Marchesiello S. Early damage detection of roller bearings using wavelet packet decomposition, ensemble empirical mode decomposition and support vector machine, Meccanica 2015; 50(3): 865-874, https://doi.org/10.1007/s11012-014-9968-z.
26. Teng W, Ding X and Zhang X et al. Multi-fault detection and failure analysis of wind turbine gearbox using complex wavelet transform, Renewable Energy 2016; 93: 591-598, https://doi.org/10.1016/j.renene.2016.03.025.

27. Wang Q, Chen H, Chen X, Yang H, Wang G. Early fault detection of gearbox using weak vibration signals, Eksploatacja i Niezawodnosc-Maintenance and Reliability 2011; 1(49): 11-15.
28. Wang Y, He Z, Zi Y. Enhancement of signal denoising and multiple fault signatures detecting in rotating machinery using dual-tree complex wavelet transform, Mechanical Systems and Signal Processing 2010; 24 (1): 119-137, https://doi.org/10.1016/j.ymssp.2009.06.015.
29. Wu Z, Huang N E. Ensemble empicrical mode decomposition: a noise-assisted data analysis method, Advances in Adaptive Data Analysis 2011; 1(01): 1-41, https://doi.org/10.1142/S1793536909000047.
30. Yu D, Wang M, Cheng X. A method for the compound fault diagnosis of gearboxes based on morphological component analysis, Measurement 2016; 91: 519-531, https://doi.org/10.1016/j.measurement.2016.05.087.
31. Zibulevsky Michael, Zeevi Yehoshua Y. Extraction of a source from multichannel data using sparse decomposition, Neurocomputing 2002; 49(1-4): 163-173, https://doi.org/10.1016/S0925-2312(02)00515-5.
32. Zosso D, Dragomiretskiy K. Variational Mode Decomposition. IEEE Transactions on Signal Processing 2014; 62(3): 531-544, https://doi.org/10.1109/TSP.2013.2288675.
33. Zuber N, Bajric R, Sostakov R. Gearbox faults identification using vibration signal analysis and artificial intelligence methods. Eksploatacja i Niezawodnosc-Maintenance and Reliability 2014; 16: 61-65.

Xin ZHANG
Jianmin ZHAO
Army Engineering University, Shijiazhuang, China

E-mail: Xinzhang91@hotmail.com, jm_zhao@hotmail.com

Article citation info:
LI J, WANG Z, REN Y, YANG D, LV X. A novel reliability estimation method of multi-state system based on structure learning algorithm. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 170–178, http://dx.doi.org/10.17531/ein.2020.1.20.

Zhifeng LI
Zili WANG
Yi REN
Dezhen YANG
Xing LV

A NOVEL RELIABILITY ESTIMATION METHOD OF MULTI-STATE SYSTEM BASED ON STRUCTURE LEARNING ALGORITHM

NOWATORSKA METODA OCENY NIEZAWODNOŚCI SYSTEMÓW WIELOSTANOWYCH W OPARCIU O ALGORYTM UCZENIA STRUKTURY

Traditional reliability models, such as fault tree analysis (FTA) and reliability block diagram (RBD), are typically constructed with reference to the function principle graph that is produced by system engineers, which requires substantial time and effort. In addition, the quality and correctness of the models depend on the ability and experience of the engineers and the models are difficult to verify. With the development of data acquisition, data mining and system modeling techniques, the operational data of a complex system considering multi-state, dependent behavior can be obtained and analyzed automatically. In this paper, we present a method that is based on the K2 algorithm for establishing a Bayesian network (BN) for estimating the reliability of a multi-state system with dependent behavior. Facilitated by BN tools, the reliability modeling and the reliability estimation can be conducted automatically. An illustrative example is used to demonstrate the performance of the method.

Keywords: *reliability analysis, Bayesian network, structure learning, multi-state system (MSS), dependent failure.*

Tradycyjne modele niezawodności, takie jak analiza drzewa błędów (FTA) czy schemat blokowy niezawodności (RBD), buduje się zazwyczaj w oparciu o tworzone przez inżynierów systemowych schematy zasad działania systemu, których przygotowanie wymaga dużych nakładów czasu i pracy. Jakość i poprawność tych modeli zależy od umiejętności i doświadczenia inżynierów, a same modele są trudne do zweryfikowania. Dzięki rozwojowi technik akwizycji i eksploracji danych oraz modelowania systemów, dane operacyjne złożonego systemu uwzględniające jego zależne, wielostanowe zachowania mogą być pozyskiwane i analizowane automatycznie. W artykule przedstawiono metodę konstrukcji sieci bayesowskiej (BN) opartą na algorytmie K2, która pozwala na ocenę niezawodności systemu wielostanowego o zachowaniach zależnych. Dzięki narzędziom BN, modelowanie i szacowanie niezawodności może odbywać się automatycznie. Działanie omawianej metody zilustrowano na podstawie przykładu.

Słowa kluczowe: *analiza niezawodności, sieć bayesowska, uczenie struktury, system wielostanowy, uszkodzenie zależne.*

1. Introduction

For estimating the reliability of a complex system, constructing an accurate reliability model of the system is essential. A variety of popular reliability models are available, such as fault trees, reliability block diagrams, and Bayesian networks [28]. Unfortunately, they require professional knowledge and experience in modeling, along with a detailed understanding of the system structure and of how the system operates. Moreover, substantial effort is required for constructing these models for complex systems, even if the systems are medium-scale. Two main difficulties are encountered in building these models. (1) Modern systems consist of hardware and software with complex interactions, which are becoming increasingly difficult to model. (2) Reliability models and function principle graphs describe different aspects of the system. Although the designers know how the systems work, they lack the skills and experience in reliability modeling and analysis. There is a gap between system functional models and reliability models.

To overcome these problems, researchers proposed automatic transformation methods for converting the function principle graphs to fault trees [1, 20] or RBDs [11]. In addition, Bucci attempted to construct dynamic fault tree and event tree from corresponding Markov models such that time-dependent failure can be considered [2]. Moreover, the dynamic reliability models such as the dynamic fault tree (DFT) or dynamic reliability block diagram can be transformed into a dynamic Bayesian network to estimate the reliability of a dynamic system [17, 19, 21]. Even if reliability models could be automatically generated from function principle graphs, they would not reflect the changes that occur when the system is operating. In addition, when human factors are involved, it is difficult to incorporate these factors into the reliability models.

Methods such as the GO methodology and the Altarica project have been proposed for overcoming the problem of models not matching the specifications of the systems under study. The GO methodology uses a straightforward inductive logic to construct system models [18]. Enhanced methods improve the performance of the GO methodology and provide additional analysis results about the systems. A new quantification algorithm that is based on BDD is proposed in [5]. The quality analysis method, in combination with FMEA, is discussed in [12] and the dynamic behavior can be modeled via the extended GO methodology [25]. Altarica is a high-level modeling language

that can describe the hierarchies of the system and the behaviors of components of the system [7, 23]. The model can be compiled into lower-level formalisms for analysis of the reliability and dependability of the systems. The methods that are discussed above can provide models that are similar to the function principle graph; however, these methods are diffcult to use in practice due to their complexity.

Data-driven system reliability estimation methods such as the prognostics and health management (PHM) method were proposed many years ago. However, the traditional PHM methods are designed for specified operational conditions and the analysis results are limited to the systems that are in operation [30]. Hence, researchers studied how to obtain a generic model from data. Zaitseva [31] presented an approach that is based on a decision tree and learns the structure function of the system that represents the states of the system and components from the source data. The method is used not only to construct reliability models of the general system but also to estimate the reliability of the corresponding human factor system [13]. Zaitseva extended this method to multi-state systems [32]. However, the approach is only suitable for relatively simple systems that have no complicated interactions among components. Doguc [4] proposed a structure learning method that is based on Bayesian networks for constructing the reliability model. However, he only utilized a binary system and did not consider the interactive effects between components.

Dependent failure is an important behavior that can substantially affect the reliability of a system. Two types of dependent failure were identified in [27]: The first type can be described by the functional-dependent gate in DFT. The second type is found in multi-state systems. Thus, dependent failure should be considered in reliability models. Unfortunately, it is difficult to identify the dependencies between components and the interactive effects between components and subsystems.

In this paper, we propose a method that is based on the structure learning algorithm for modeling and estimating the reliability of multi-state systems while considering dependent failures. We focus on the dependency behavior between components.

The remainder of this paper is organized as follows: Section 2 briefly summarizes the multi-state system with correlative behavior. Section 3 presents a methodology for modeling the cause and effect relationships of multi-state systems that is based on the K2 algorithm [3] and evaluating conditional probability tables (CPTs). In Section 4, an example is used to demonstrate the entire modeling process. Section 5 presents the experimental analysis results on the accuracy and performance of the methodology. Finally, in Section 6, the conclusions of this work are presented and discussed.

2. Multi-state system with dependency behavior

Multi-state systems were introduced in 1968. Many researchers gradually contributed to the reliability theory of multi-state systems and developed a variety of methods for evaluating the reliability of multi-state systems. The related works are referenced in [14, 29].

For simplicity, traditional reliability modeling methods assume that the components are independent. However, this assumption is not practical for real engineering systems. In [27], the authors described several dependent failure scenarios, such as common cause failure, load-sharing, cascade failure [24], sequential failure, cross-system dependencies and interaction between components. These dependent failures would severely affect the reliability estimation of the systems.

Thus, researchers improved the traditional models to include dependent failures. A stochastic process [16] is an intuitive model for representing the correlations between components. Song [26] presented stochastic multivalued models for evaluating the reliability of an MSS with dependent multistate components. However, the space explosion problem poses substantial challenges for large-scale multi-state systems. So researchers explored the combinatorial methods to model the multi-state system with dependent failure. Levitin [15] extended the universal generating function approach to multi-state systems with dependent elements. The dependent failure is of the second type [26]. Nagayama [22] analyzed the reliability of a multi-state system with partially dependent components based on the multi-valued decision diagrams.

In the works that are discussed above, the dependencies between components are formulated as conditional probabilities. The problem can be more easily represented by a Bayesian network, which can describe complicated relationships. Thus, BN is an alternative way of modeling a multi-state system with dependency behavior. Various BN tools such as BNT [6] support BN inference and implement many learning algorithms. Consequently, BN is a more suitable method for modeling multi-state systems with consideration of the dependency behavior.

3. Methodology

In this section, a reliability estimation method for multi-state systems that is based on the K2 algorithm [3] and the point estimation method is proposed. The dependent relationships among components, subsystems and the system are learned from data via the K2 algorithm. The parameters of the BN are estimated via the point estimation method. Then, the reliability of the multi-state system with consideration of the customer demand can be evaluated. Facilitated by BN tools, the whole process can be conducted automatically.

3.1. Structure learning algorithm K2

The K2 algorithm was proposed in [3] as a heuristic-search method. This algorithm assumes the following: 1) the variables are ordered and 2) all structures are equally likely. According to the assumptions, high-rank variables will not be the parents of low-rank variables. Hence, the search space of the parent sets of a variable can be reduced substantially.

The K2 algorithm consists of two main components:

1) A scoring function that quantifies the associations and ranks the parent sets according to their scores:

$$g(i,\pi_i)=\prod_{j=1}^{q_i}\frac{(r_i-1)!}{(N_{ij}+r_i-1)!}\prod_{k=1}^{r_i}N_{ijk}! \tag{1}$$

i: index of the components.

π_i: set of parents of components x_i.

q_i: $|\phi_i|$.

ϕ_i: list of all possible instantiations of the parents of x_i in database D, namely, if $x_1,\ldots,x_s$ are the parents of x_i, then ϕ_i is the Cartesian product $\{x_{10},x_{11},\cdots,x_{1j},\cdots x_{1n_1}\}\times\cdots\times\{x_{s0},x_{s1},\cdots,x_{sj},\cdots x_{sn_s}\}$. n_1, n_s represent the numbers of the states of the components.

$r_i=|V_i|$.

V_i: list of all possible state values of the components x_i.

α_{ijk}: the number of cases in D in which component x_i is instantiated with its kth value and the parents of x_i in π_i are instantiated with the jth instantiation in ϕ_i.

$N_{ij}=\sum_{k=1}^{r_i}\alpha_{ijk}$: the number of instances in the database in which the parents of x_i in π_i are instantiated with the jth instantiation in ϕ_i.

To improve the run-time speed of K2, the logarithmic version of the equation above is implemented in this paper. The equation is described as follows:

$$\begin{aligned}\log(g(i,\pi_i)) &= \sum_{j=1}^{q_i}\log\left(\frac{(r_i-1)!}{(N_{ij}+r_i-1)!}\right)+\sum_{k=1}^{r_i}\log(N_{ijk}!)\\ &= \sum_{j=1}^{q_i}\Big[\log((r_i-1)!)-\log((N_{ij}+r_i-1)!)\Big]+\sum_{k=1}^{r_i}\log(N_{ijk}!)\\ &= \sum_{j=1}^{q_i}\Big[\log(\Gamma(r_i-1))-\log(\Gamma(N_{ij}+r_i-1))\Big]+\sum_{k=1}^{r_i}\log(\Gamma(N_{ijk}))\end{aligned} \tag{2}$$

where $\Gamma(\cdot)$ denotes the gamma function and the other parameters are as defined above;

2) A greedy-search method that incrementally adds nodes to the parent set to reduce the search space.

With the heuristic, the K2 algorithm does not need to consider all possible parents sets; it adds incrementally the parent whose addition most increases the probability of the resulting structure. If the addition of no single parent can increase the probability, no additional parents are added to the node: Initially, node x_i has no parents and the nodes $x_1, x_2, \cdots, x_{i-1}$ are candidates of the parent sets. The parent sets may be $\phi, \{x_1\}, \{x_2\}, \cdots, \{x_{i-1}\}, \{x_1, x_2\}, \cdots, \{x_1, x_{i-1}\}, \cdots, \{x_1, x_2, \cdots, x_{i-1}\}$. Hence, the parent set space is reduced substantially. In addition, to increase the efficiency of the algorithm, the K2 algorithm uses a parameter u to restrict the maximum number of parents. The pseudocode of the K2 algorithm can be found in [3].

According to the description above, to use the K2 algorithm to learn the exact Bayesian network, the variables' ordering should correspond to the practical operational mode of the system. For example, a subsystem should be higher in the ordering than its child nodes. In addition, the parameter u should be set reasonably to balance the efficiency and the correctness of the Bayesian network. In [3], several suggestions are proposed for obtaining the most probable structure.

3.2. Probability distribution estimation

Via the K2 algorithm, the relationships among the components, subsystems and system can be obtained. To evaluate the reliability of MSS, the parameters of the Bayesian network should be estimated using the data.

Traditional reliability methods typically assume that the failure times of the components follow a probability distribution. However, this assumption may not hold in practice. Researchers considered integrating data from various levels of the system to reduce the uncertainty in the system reliability assessment. These methods [8, 9] can use the data to estimate the parameters of multi-state components. However, these approaches require prior knowledge about the system structure and about the cause and effect between components. These conditions are difficult to satisfy when we only have the function principle graph and the operation data of the system. Consequently, statistical estimation theory is used to determine the probability distribution.

There are two main types of estimation procedures in statistics: point and interval estimation. For convenience, point estimation is applied in this paper. Interval estimation is also applicable in our method.

Table 1. Probability distribution of component

X_i	1	2	$\cdots$	j	$\cdots$	n_i
P	$\frac{H_{i1}}{H}$	$\frac{H_{i2}}{H}$	$\cdots$	$\frac{H_{ij}}{H}$	$\cdots$	$\frac{H_{in_i}}{H}$

First, according to the BN model that was learned from the data, the nodes without parents can be identified. The probability distributions of these nodes can be calculated easily. For node X_i, denotes the number of times that state j of component X_i occurs in the data as H_{ij} and the number of all instances in the data as H_i. The state probability distribution of X_i is presented in Table 1, where n_i denotes the number of states of component X_i.

Second, the conditional probability distributions of the nodes that have parents are estimated. For example, the structure of nodes X_i, X_j, X_k is illustrated in Fig. 1.

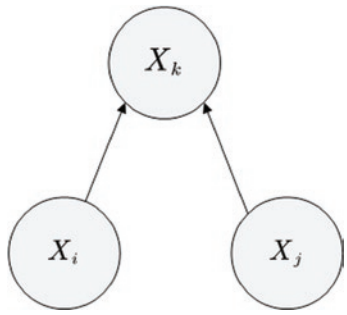


Fig. 1. Illustrative example

The CPT of X_k can be estimated via the following formula:

$$\begin{aligned}&P\{X_k = c_{kr} \mid X_i = c_{iw}, X_j = c_{jv}\}\\ &= \frac{P\{X_k = c_{kr}, X_i = c_{iw}, X_j = c_{jv}\}}{P\{X_i = c_{iw}, X_j = c_{jv}\}}\\ &= \frac{H\{X_k = c_{kr}, X_i = c_{iw}, X_j = c_{jv}\}}{H\{X_i = c_{iw}, X_j = c_{jv}\}}\end{aligned} \tag{3}$$

$H\{X_k = c_{kr}, X_i = c_{iw}, X_j = c_{jv}\}$ denotes the number of instances when $X_k = c_{kr}, X_i = c_{iw}, X_j = c_{jv}$. $H\{X_i = c_{iw}, X_j = c_{jv}\}$ is also available.

According to the method above, a BN that represents the dependent relationships and all parameters that represent the logical relationships can be obtained. Hence, the reliability of the MSS can be evaluated. Assume the system has M unique states. The reliability of the multi-state system can be expressed by the following formula:

$$R(t) = \sum_{s_i \geq w(\mathrm{t})}^{M} P\{S(\mathrm{t}) = s_i\} \tag{4}$$

where $S(\mathrm{t})$ denotes the current state of the system and $w(\mathrm{t})$ denotes the demand for the system at time t. According to the model that is obtained above, the joint probability distribution $P(S, X_1, \cdots, X_j, \cdots, X_m)$ can be calculated. Thus, the marginal probability distribution $P\{S(\mathrm{t}) = s_i\} = p_{s_i}$ is also obtained.

4. Illustrative example

4.1. Classification of Failure Conditions

In this paper, the task processing system example from [15] is used to demonstrate our method. The system logic diagram is shown as Fig. 2.

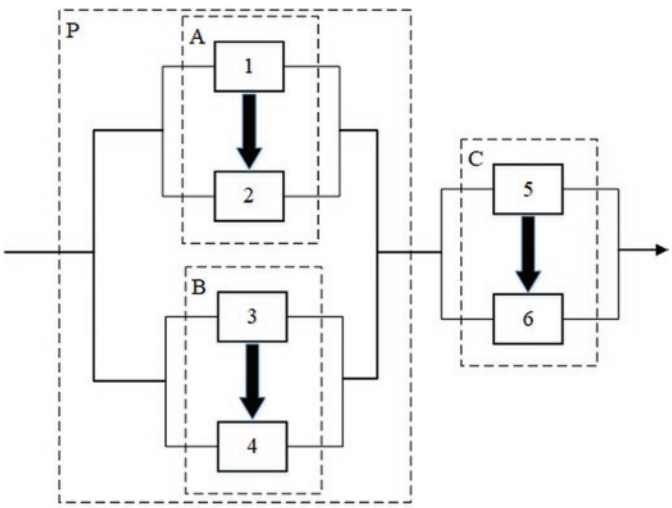


Fig. 2. System logic diagram of the task processing system

The system consists of three independent computing blocks: A, B and C. Blocks A and B are constructed in parallel. Then, the parallel structure P and block C are arranged in series. Each block is composed of two processing units that differ in terms of priority. Elements 1, 3 and 5 have high priorities in blocks A, B and C. When the high-priority unit accesses a database, the low priority unit must to wait for the operation to be completed. Thus, the processing speed of the low-priority unit is affected by the load of the high-priority unit. The performance distributions of elements 1, 3 and 5 and the conditional performance distributions of elements 2, 4 and 6 can be found in [15].

The Bayesian network model of the task processing system is illustrated in Fig. 3. Nodes G_1, G_2, G_3, G_4, G_5, and G_6 represent processing units 1, 2, 3, 4, 5, and 6. Nodes A, B and C represent computing blocks A, B and C. Node P represents the subsystem that consists of blocks A and B. Node S represents the task processing system.

The BN is used to evaluate the performance in model learning from the data. In the next section, the logic sampling method [10] is used to randomly generate virtual instances to demonstrate the relationship between accuracy of the BN and the number of observations.

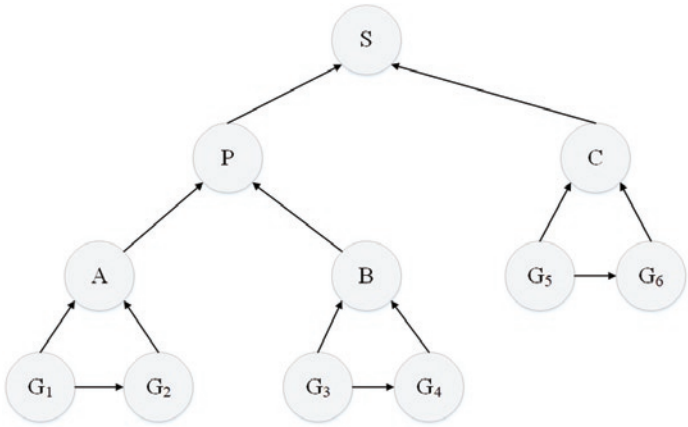


Fig. 3. Bayesian network of the task processing system

The performance of each element is discrete; hence, we treat the performance as the state of the element, namely, the performance and the state of the element have the same meaning in this paper.

The performance distributions of elements 1, 3 and 5 are presented in Table 2. The conditional performance distributions of elements 2, 4 and 6 are presented in Table 3, Table 4 and Table 5.

The conditional performance distributions of block A are presented in Table 6. Similarly, the conditional performance distributions of nodes B, C, P and S can be obtained.

The data obtained based on the monitoring and presented in Table 7, which collected 200 samples of the task processing system. The data set demonstrates some combinations of component states and the corresponding performance levels of the system. Though the data set is small, the relationships among the components and system can be inferred partially. The detail process is illustrated in the next subsection.

Table 2. Performance distributions of elements 1, 3 and 5

G_1	0	10	20	30	40	50
P_1	0.05	0.05	0.1	0.1	0.5	0.2
G_3	0	20	60			
P_3	0.1	0.7	0.2			
G_5	0	80	100			
P_5	0.1	0.2	0.7			

Table 3. Conditional performance distribution of element 2

G_1		0	10	20	30	40	50
G_2	0	0.05	0.05	0.05	0.05	0.1	0.1
	15	0.15	0.15	0.55	0.55	0.9	0.9
	30	0.8	0.8	0.4	0.4	0	0

Table 4. Conditional performance distribution of element 4

G_3		0	20	60
G_4	0	0.05	0.05	0.05
	15	0.15	0.35	0.95
	30	0.8	0.6	0

Table 5. Conditional performance distribution of element 6

G_5		0	80	100
G_6	0	0.05	0.1	0.1
	30	0.15	0.4	0.6
	50	0.8	0.5	0.3

4.2. Reliability estimation of the task processing system

According to the system logic diagram in Fig. 2, the ranking of the nodes depends on the hierarchy of the nodes, which will affect the efficiency of the K2 algorithm. The ranks and the number of performance value of the components are listed in Table 10.

Using the K2 algorithm, the associations among the components, subsystems and system can be identified and the score function values that correspond to parents can be obtained. The results are listed in

Table 6. Conditional probability distribution of block A

G_1	0			10			20			30			40			50		
G_2	0	15	30	0	15	30	0	15	30	0	15	30	0	15	30	0	15	30
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0
45	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
55	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Table 7. Data obtained based on the monitoring of the task processing system

No.	G_1	G_2	A	G_3	G_4	B	P	G_5	G_6	C	S
1	40	15	55	60	15	0	55	100	30	130	55
2	40	15	55	20	30	50	105	0	50	50	50
3	50	15	65	20	15	35	100	0	50	50	50
4	20	30	50	0	30	30	80	100	30	130	80
5	40	15	55	60	15	75	115	80	50	130	115
6	10	30	40	20	30	50	90	80	30	110	80
7	20	15	35	0	30	30	65	80	50	130	65
8	10	30	40	20	30	50	90	80	30	110	90
9	0	15	15	20	30	50	65	100	0	100	65
10	0	0	0	20	30	50	50	0	50	50	50
...	...	...	...	...	...	...	...	...	...	...	...
200	30	30	60	20	30	50	110	100	50	150	110

Table 10. Ranks and the number of performance value of the components

Rank	Component/Subsystem	The number of performance value
1	G_1	6
2	G_2	3
3	A	15
4	G_3	3
5	G_4	3
6	B	9
7	P	33
8	G_5	3
9	G_6	3
10	C	8
11	S	30

Table 11. Associations and score function values

Component/Sub-system	Parent set	Score function value
G_1	ϕ	−292.63
G_2	$\{G_1\}$	−112.85
A	$\{G_1, G_2\}$	−184.89
G_3	ϕ	−165.17
G_4	$\{G_3\}$	−142.15
B	$\{G_3, G_4\}$	−93.49
G_5	ϕ	−173.42
G_6	$\{G_5\}$	−191.04
C	$\{G_5, G_6\}$	−107.88
P	$\{A, B\}$	−427.34
S	$\{P\}$	−429.06

Table 11 and the raw system structure inferred from the data is shown in Fig. 4. The system structure lacks of the relationship between component C and the system and the reason is that the data in Table 7 just represents partial state combinations of the system. The correlation of accuracy and data volume is analyzed in the next section.

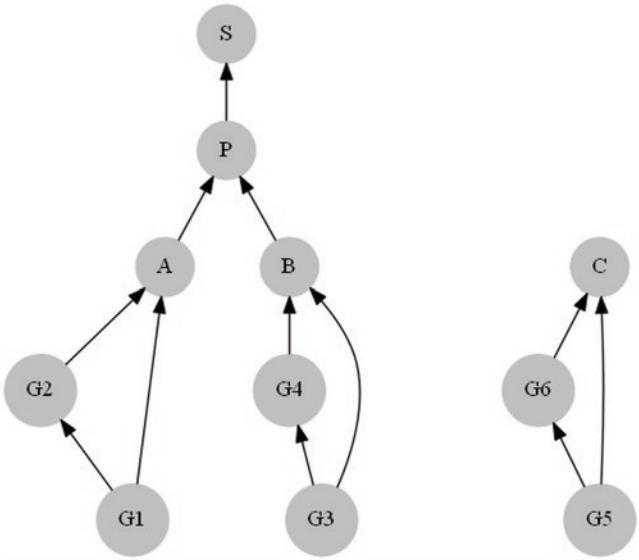


Fig. 4. The system structure inferred from the data

Table 12. Probability estimator of component G_1

G_1	0	10	20	30	40	50
p	0.04	0.05	0.06	0.11	0.515	0.225

Table 13. Conditional probability distribution estimator of component G_2

G_1		0	10	20	30	40	50
G_2	0	0	0	0	0	0.1165	0.1111
	15	0.25	0.2	0.25	0.5	0.8835	0.8889
	30	0.75	0.8	0.75	0.5	0	0

The performance distribution parameters of the components or subsystems can be estimated by the data obtained based on the monitoring. The approximate performance distributions of elements G_1 and G_2 in the data are presented in Table 12 and Table 13. The performance distributions for the other elements can be obtained via the same approach.

Then, the reliability of the task processing system with consideration of the demand can be evaluated and the system reliability as a function of the demand is plotted in Fig. 5. The result is compared with the reliability that was calculated via the UGF method in Fig. 5.

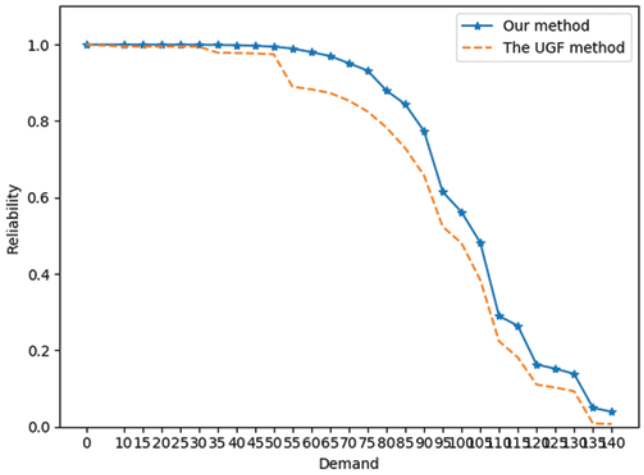


Fig. 5. Reliability comparison between our method and the UGF method

5. Experimental analysis

5.1. Generating random instances

The logic sampling generates an instance by randomly selecting values from the probability tables or conditional probability tables. The nodes are visited from the root nodes to the leaves; hence, all nodes in the BN can be instantiated once all nodes have been traversed. The detailed algorithm is available in [10].

Next, let us work through one round of simulation for the task processing system. Consider the root node G_1 of this network as an example. The random number generator produces a value between 0 and 1. Assume that the random number is 0.56. Then, the corresponding performance of unit 1 is 40, as illustrated in Fig. 6.

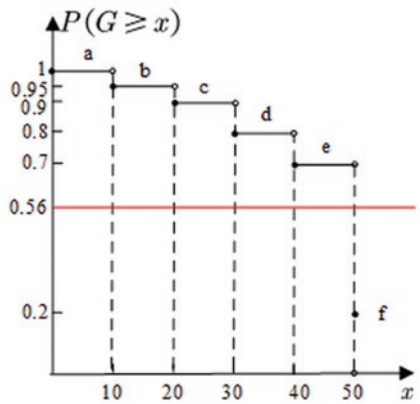


Fig. 6. Cumulative performance curve for processing unit 1

Table 7. Performance values of all root nodes

G_1	G_3	G_5
40	20	80

Table 8. Performance instances of units 2, 4 and 6

G_2	G_4	G_6
15	30	30

Table 9. Full combinations of values in this simulation round

G_1	G_2	G_3	G_4	G_5	G_6	A	B	C	P	S
40	15	20	30	80	30	55	50	110	105	105

The performance values of other root nodes can be obtained as well. The performance instances of all root nodes are listed in Table 7.

Next, the value for child node G_2 should be generated. According to Table 3, the conditional cumulative performance curve for processing unit 2 can be calculated. The random number is 0.72; hence, the performance value of unit 2 is 15. The performance instances of units 2, 4 and 6 are listed in Table 8.

Then, the performance values of other nodes in the BN can be computed via the method that is demonstrated above. Finally, the full combinations of values in this simulation round are listed in Table 9.

The simulation is repeated 20000 times and 20000 sets of cases are generated as the operation data of the system. These cases form the training data set.

5.2. The correlation between accuracy and data volume

The number of observations that are used to discover the associations among nodes affects the efficiency of the K2 algorithm and the accuracy of the constructed BN. The error rate that is defined in [4] was used here to analyze the relationship between accuracy of the BN and the number of observations. The error rate can be calculated as follows:

$$\rho = \frac{A_{FP} + A_{FN}}{A_T} \tag{5}$$

A_{FP} : the number of associations that do not exist in the actual BN.

A_{FN} : the number of associations in the actual BN that are missed.

A_T : the number of the associations in the constructed BN.

As the amount of data increases, structure learned from data approximates the actual BN with increasing accuracy. The BN models that correspond to various numbers of instances are illustrated in Fig. 6.

The last Bayesian network has same structure as the model that is illustrated in Fig. 3. The BN that is constructed using a dataset with 1000 instances is similar to the actual BN of the task processing system. The BN that was built with 2500 instances has the same associations as the BN with 1000 instances. Hence, the additional data do not provide more information regarding the associations between block C and the system. With the increase in the amount of data, node G_5 becomes associated with node S. However, the performance of the system depends on block C and not on element 5. Although the additional data provide new information about the relation between the system and block C, the associations do not represent the real relationship. Moreover, the error rates of the BNs with 3000 and 12000 instances become higher than that of the BN with 2500 instances. In the course of training, the error rate may fluctuate with the increase in the amount of data before finally converging. The error rate curve is plotted in Fig. 7

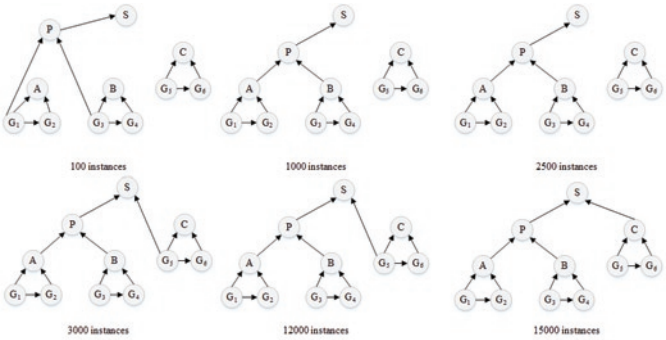


Fig. 7. BN models for various numbers of instances

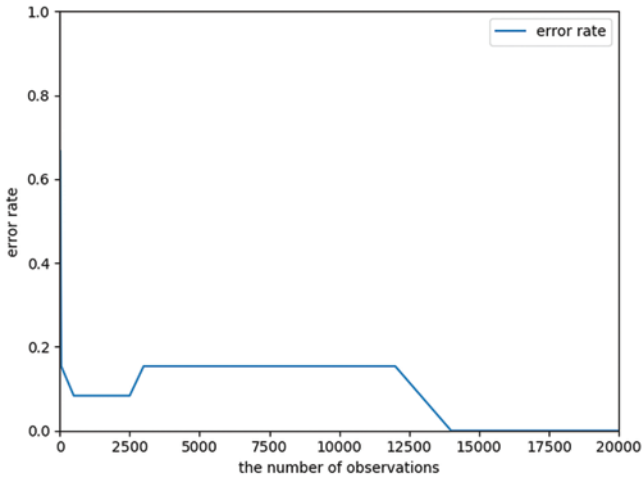


Fig. 8. Error rate versus the number of observations

According to Fig. 6 and Fig. 7, the dependent relationships of computing blocks A and B and the system are easily constructed; however, more data are required for accurately establishing the associations of computing block C and the system. That is because the task processing system is a series system and its performance depends on the minimal performance of subsystem P and block C. Hence, the accuracy may vary with the structure of the system.

6. Conclusions and Discussion

Traditional reliability methods strongly depend on the ability and experience of the engineers, which results in differences among reliability models that model the same system. With the development of data acquisition and data mining techniques, the operational data of

the systems can be monitored and analyzed for reliability estimation and system optimization. In this paper, a new method that is based on the K2 algorithm is proposed for constructing a reliability model of a system and for estimating the parameters of components from data. In the illustrative example, the Bayesian network model that is learned from the data is the same as the Bayesian network model that we constructed. Moreover, comparing with the universal generating function method, the results of the two methods are very close. Hence, this approach is effective.

According to experimental results, the efficiency of the method may depend on the structure of the systems. Thus, determining how to integrate the structure of the systems as prior knowledge into the process of structure learning will constitute our future work. This paper does not consider the scenario of missing data, which is common in practice. Determining how to handle missing data will be left for future work.

In conclusion, this method is suitable for reliability estimation of complex systems while considering multi-state and dependency behavior. Facilitated by BN tools, reliability modeling and reliability estimation can be conducted automatically without human intervention.

Acknowledgement

The research work financed with the means of National Natural Science Foundation of China (Grant No. 51805018).

References

1. Bhagavatula A, Tao J, Dunnett S, et al. A new methodology for automatic fault tree construction based on component and mark libraries. Safety and Reliability 2016; 36(2): 62-76, https://doi.org/10.1080/09617353.2016.1219934.
2. Bucci P, et al. Construction of event-tree/fault-tree models from a Markov approach to dynamic system reliability. Reliability Engineering & System Safety 2008; 93(11): 1616-1627, https://doi.org/10.1016/j.ress.2008.01.008.
3. Cooper G F, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. Machine Learning 1992; 9(4): 309-347, https://doi.org/10.1007/BF00994110.
4. Doguc O, Ramirez-Marquez J E. A generic method for estimating system reliability using Bayesian networks. Reliability Engineering & System Safety 2009; 94(2): 542-550, https://doi.org/10.1016/j.ress.2008.06.009.
5. Fan J, Ren Y, Liu L L. A new GO methodology algorithm based on BDD. Advanced Materials Research 2013; 791-793:1134-1138, https://doi.org/10.4028/www.scientific.net/AMR.791-793.1134.
6. François, Olivier & Leray, Philippe. BNT Structure Learning Package: Documentation and Experiments. 2008.
7. G. Point and A. Rauzy. AltaRica: Constraint automata as a description language. European Journal on Automation 1999; 33(8-9): 1033-1052.
8. Hao Z, Guo J, Zeng S. Fully Bayesian reliability assessment of multi-state systems with overlapping data. Journal of Systems Engineering and Electronics 2017; 28(1): 187-198, https://doi.org/10.21629/JSEE.2017.01.21.
9. Jackson C, Mosleh A. Bayesian inference with overlapping data: Reliability estimation of multi-state on-demand continuous life metric systems with uncertain evidence. Reliability Engineering & System Safety 2016; 145: 124-135, https://doi.org/10.1016/j.ress.2015.09.006.
10. Jensen F V. Bayesian Artificial Intelligence, second edition. Chapman & Hall/crc Boca Raton Fl 2010.
11. Liu X , Ren Y , Wang Z , et al. Modeling method of SysML-based reliability block diagram. Proceedings 2013 International Conference on Mechatronic Sciences 2013: 206-209, https://doi.org/10.1109/MEC.2013.6885073.
12. Liu L, Fan D, Wang Z, et al. Enhanced GO methodology to support failure mode, effects and criticality analysis. Journal of Intelligent Manufacturing 2019; 30(2): 1451-1468, https://doi.org/10.1007/s10845-017-1336-0.
13. Levashenko V , Zaitseva E , Kvassay M , et al. Reliability estimation of healthcare systems using Fuzzy Decision Trees. 2016 Federated Conference on Computer Science and Information Systems 2016: 331-340, https://doi.org/10.15439/2016F150.
14. Levitin, Gregory & Xing, Liudong. Multi-state systems. Reliability Engineering & System Safety 2017; 166: 1-2, https://doi.org/10.1016/j.ress.2017.06.008.
15. Levitin G. A universal generating function approach for the analysis of multi-state systems with dependent elements. Reliability Engineering & System Safety 2004; 84(3): 285-292, https://doi.org/10.1016/j.ress.2003.12.002.
16. Lesanovsky, A. Multistate Markov models for systems with dependent units. IEEE Transactions on Reliability 1988; 37(5): 505-511, https://doi.org/10.1109/24.9872.
17. Li K, Yi R, Ma Z. Reliability analysis of dynamic reliability blocks through conversion into dynamic bayesian networks. 2016 IEEE International Conference on Industrial Engineering and Engineering Management 2016: 1330-1334, https://doi.org/10.1109/IEEM.2016.7798094.
18. Matsuoka T, Kobayashi M, Takemura K. The GO-FLOW Methodology: A Reliability Analysis of the Emergency Core Cooling System of a Marine Reactor Under Accident Conditions. Nuclear Technology 1989; 84(3): 285-295, https://doi.org/10.13182/NT89-A34212.
19. Mi J, Li Y, Huang H Z, et al. Reliability analysis of multi-state systems with common cause failure based on Bayesian Networks. Eksploatacja i Niezawodnosc – Maintenance and Reliability.2013; 15(2): 169-175, https://doi.org/10.1109/ICQR2MSE.2012.6246417
20. Majdara A, Wakabayashi T, Soares C G. Component-based modeling of systems for automated fault tree generation. Reliability Engineering & System Safety 2009; 94(6): 1076-1086, https://doi.org/10.1016/j.ress.2008.12.003.
21. Montani S, Portinale L, Bobbio A , et al. RADYBAN : A tool for reliability analysis of dynamic fault trees through conversion into dynamic Bayesian networks. Reliability Engineering & System Safety 2008; 93(7), 922–932, https://doi.org/10.1016/j.ress.2007.03.013.
22. Nagayama, Shinobu & Sasao, Tsutomu & Butler, Jon & Thornton, Mitch & Manikas, Theodore. Analysis Methods of Multi-state Systems Partially Having Dependent Components Using Multiple-Valued Decision Diagrams. 2014 IEEE 44th International Symposium on Multiple-Valued Logic 2014: 190-195. https://doi.org/10.1109/ISMVL.2014.41.
23. Prosvirnova T. The AltaRica 3.0 Project for Model-Based Safety Assessment. 2013 11th IEEE International Conference on Industrial Informatics (INDIN) 2013: 741-746, https://doi.org/10.1109/INDIN.2013.6622976.
24. Peng R. Reliability of interdependent networks with cascading failures. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 273–277, http://doi.org/10.17531/ein.2018.2.13.
25. Ren Y, Fan, Dongming, Wang, Zili, et al. System Dynamic Behavior Modelling based on Extended GO Methodology. IEEE Access 2018; 6: 22513-22523, http://doi.org/10.1109/ACCESS.2018.2816165.

26. Song X , Zhai Z , Liu Y , et al. A Stochastic Approach for the Reliability Evaluation of Multi-State Systems with Dependent Components. Reliability Engineering & System Safety 2018; 170: 257-266, https://doi.org/10.1016/j.ress.2017.10.015.
27. Sun Y, Ma L, Mathew J, et al. An analytical model for interactive failures. Reliability Engineering & System Safety 2006; 91(5): 495-504, https://doi.org/10.1016/j.ress.2005.03.014.
28. WANG H, DUAN F, MA J. Reliability analysis of complex uncertainty multi-state system based on Bayesian network. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21(3): 419–429, http://dx.doi.org/10.17531/ein.2019.3.8.
29. Yingkui G , Jing L . Multi-State System Reliability: A New and Systematic Review. Procedia Engineering 2012; 29: 531-536, https://doi.org/10.1016/j.proeng.2011.12.756.
30. Zio, Enrico. Some Challenges and Opportunities in Reliability Engineering. IEEE Transactions on Reliability 2016; 65(4): 1769-1782, http://dx.doi.org/10.1109/TR.2016.2591504.
31. Zaitseva E , Levashenko V . Construction of a Reliability Structure Function Based on Uncertain Data. IEEE Transactions on Reliability 2016; 65 (4): 1710-1723, http://dx.doi.org/10.1109/TR.2016.2578948.
32. Zaitseva, E., Levashenko, V., Kvassay, M., & Rabcan, J. Application of Ordered Fuzzy Decision Trees in Construction of Structure Function of Multi-State System. In International Conference on Information and Communication Technologies in Education, Research, and Industrial Applications 2016: 56–75, http://dx.doi.org/10.1007/978-3-319-69965-3_4.

Zhifeng LI
Zili WANG
Yi REN
Dezhen YANG
Xing LV
School of Reliability and Systems Engineering
Beihang University
No.37 Xueyuan RD. Haidian, 100191, Beijing, China

E-mails: lizhifeng@buaa.edu.cn, wangzili@buaa.edu.cn, renyi@buaa.edu.cn, dezhenyang@buaa.edu.cn, lvxing@buaa.edu.cn

Editorial Board

Prof. Andrzej Niewczas
Chair of Editorial Board
President of the Board of the Polish Maintenance Society

Prof. Holm Altenbach
Otto-von-Guericke-Universität, Magdeburg, Germany

Prof. John Andrews
University of Nottingham, Nottingham, UK

Prof. Karol Andrzejczak
Poznań University of Technology, Poznań, Poland

Prof. Christophe Bérenguer
Institut Polytechnique de Grenoble, Grenoble, France

Prof. Gintautas Bureika
Vilnius Gediminas Technical University, Vilnius, Lithuania

Dr Alireza Daneshkhah
Warwick Centre for Predictive Modelling
University of Warwick, UK

Prof. Luis Andrade Ferreira
University of Porto, Porto, Portugal

Prof. Mitra Fouladirad
Troyes University of Technology, France

Dr Ilia Frenkel
Shamoon College of Engineering, Beer Sheva, Israel

Prof. Olgierd Hryniewicz
Systems Research Institute of the Polish Academy of Science, Warsaw, Poland

Prof. Hong-Zhong Huang
University of Electronic Science and Technology of China, Chengdu, Sichuan, China

Prof. Marianna Jacyna
Warsaw University of Technology, Warsaw, Poland

Prof. Vaclav Legat
Czech University of Agriculture, Prague, Czech Republic

Prof. Jerzy Merkisz
Poznań University of Technology, Poznań, Poland

Prof. Gilbert De Mey
University of Ghent, Belgium

Prof. Maria Francesca Milazzo
University of Messina, Italy

Prof. Tomasz Nowakowski
Wrocław University of Technology, Wrocław, Poland

Prof. Marek Orkisz
Rzeszów University of Technology, Rzeszów, Poland

Prof. François Pérès
Touluse University, Touluse, France

Prof. Jan Szybka
AGH University of Science and Technology, Cracow, Poland

Prof. Marcin Ślęzak
Motor Transport Institute, Warsaw, Poland

Prof. Katsumi Tanaka
Kyoto University, Kyoto, Japan

Prof. David Vališ
University of Defence, Brno, Czech Republic

Prof. Lesley Walls
University of Strathclyde, Glasgow, Scotland

Prof. Min Xie
City University of Hong Kong, Hong Kong

Prof. Irina Yatskiv
Riga Transport and Telecommunication Institute, Latvia

The Journal is indexed and abstracted in the Journal Citation Reports (JCR Science Edition), Scopus, Science Citation Index Expanded (SciSearch®) and Index Copernicus International.

The Quarterly appears on the list of journals credited with a high impact factor by the Polish Ministry of Science and Higher Education and is indexed in the Polish Technical Journal Contents database – BAZTECH and the database of the Digital Library Federation.

Ministry of Science and Higher Education
Republic of Poland

Task *„Implementation of procedures ensuring the originality of scientific papers published in the quarterly „Eksploatacja i Niezawodność – Maintenance and Reliability"* financed under contract 532/P-DUN/2018 from the funds of the Minister of Science and Higher Education for science dissemination activities.

All the scientific articles have received two positive reviews from independent reviewers.
Our 2018 Impact Factor is 1.806

Editorial staff:	Dariusz Mazurkiewicz, PhD, DSc (Eng), Associate Professor (Editor-in-Chief, Secretary of the Editorial Board) Tomasz Klepka, PhD, DSc (Eng), Associate Professor (Deputy Editor-in-Chief) Teresa Błachnio-Krolopp, MSc (Eng) (Editorial secretary) Andrzej Koma (Typesetting and text makeup) Krzysztof Olszewski, PhD (Eng) (Webmaster)
Publisher:	Polish Maintenance Society, Warsaw
Scientific patronage:	Polish Academy of Sciences Branch in Lublin
Address for correspondence:	"Eksploatacja i Niezawodność" – Editorial Office ul. Nadbystrzycka 36, 20-618 Lublin, Poland e-mail: office@ein.org.pl http://www.ein.org.pl/
Circulation:	550 copies

INFORMATION FOR AUTHORS

Eksploatacja i Niezawodnosc – Maintenance and Reliability – the journal of the Polish Maintenance Society, under the scientific supervision of the Polish Academy of Sciences (Branch in Lublin), published four times a year.

The scope of the Quarterly

The quarterly *Eksploatacja i Niezawodnosc – Maintenance and Reliability* publishes articles containing original results of experimental research on the durability and reliability of technical objects. We also accept papers presenting theoretical analyses supported by physical interpretation of causes or ones that have been verified empirically. *Eksploatacja i Niezawodność – Maintenance and Reliability* also publishes articles on innovative modeling approaches and research methods regarding the durability and reliability of objects.

The following research areas are particularly relevant to the journal:

1. degradation processes of mechanical and biomechanical systems,
2. diagnosis and prognosis of operational malfunctions and failures.
3. analysis of failure risk/wear,
4. reliability-and-environmental-safety engineering in the design, manufacturing and maintenance of objects,
5. management and rationalization of object maintenance,
6. risk management in the processes of operation and maintenance,
7. the human factor and human reliability in operation and maintenance systems.

Terms and Conditions of Publication

The quarterly *Eksploatacja i Niezawodnosc – Maintenance and Reliability* publishes only original papers written in English or in Polish with an English translation. Translation into English is done by the Authors after they have received information from the Editorial Office about the outcome of the review process and have introduced the necessary modifications in accordance with the suggestions of the referees! Acceptance of papers for publication is based on two independent reviews commissioned by the Editor. The publication fee for one text is 320 EUR (all graphics in color included).

The quarterly *Eksploatacja i Niezawodnosc – Maintenance and Reliability* proceeds entirely online at submission.ein.org.pl

Technical requirements

- After receiving positive reviews and after acceptance of the paper for publication, the text must be submitted in a Microsoft Word document format.
- Drawings and photos should be additionally submitted in the form of high resolution separate graphical files in the TIFF, SVG, AI or JPG formats.
- A manuscript should include: names of authors, title, abstract, and key words that should complement the title and abstract (in Polish and in English), the text in Polish and in English with a clear division into sections (please, do not divide words in the text); tables, drawings, graphs, and photos included in the text should have descriptive two-language captions, if this can be avoided, no formulae and symbols should be inserted into text paragraphs by means of a formula editor; references (written in accordance with the required reference format); author data – first names and surnames along with scientific titles, affiliation, address, phone number, fax, and e-mail address.

The Editor reserves the right to abridge and adjust the manuscripts. All submissions should be accompanied by a submission form.

Detailed instructions to Authors, including evaluation criteria can be found on the journal's website: www.ein.org.pl

Editor contact info

Editorial Office of „Eksploatacja i Niezawodnosc – Maintenance and Reliability"
Nadbystrzycka 36, 20-618 Lublin, Poland
e-mail: office@ein.org.pl

INFORMATION FOR SUBSCRIBERS

Fees

Yearly subscription fee (four issues) is 100 złoty and includes delivery costs. Subscribers receive any additional special issues published during their year of subscription free of charge.

Orders

Subscription orders along with authorization to issue a VAT invoice without receiver's signature should be sent to the Editor's address.

Note

In accordance with the requirements of citation databases, proper citation of publications appearing in our Quarterly should include the full name of the journal in Polish and English without Polish diacritical marks, i.e.,

Eksploatacja i Niezawodnosc – Maintenance and Reliability.

No text or photograph published in „Maintenance and Reliability" can be reproduced without the Editor's written consent.

dr inż. Jerzy Andruszkiewicz
Instytut Elektroenergetyki
Politechnika Poznańska
ul. Piotrowo 3a, 60-965 Poznań, Polska
jerzy.andruszkiewicz@put.poznan.pl

prof. dr hab. inż. Józef Lorenc
Instytut Elektroenergetyki
Politechnika Poznańska
ul. Piotrowo 3a, 60-965 Poznań, Polska
jozef.lorenc@put.poznan.pl

mgr inż. Krzysztof Łowczowski
Instytut Elektroenergetyki
Politechnika Poznańska
ul. Piotrowo 3a, 60-965 Poznań, Polska
krzysztof.lowczowski@put.poznan.pl

mgr inż. Agnieszka Weychan
Instytut Elektroenergetyki
Politechnika Poznańska
ul. Piotrowo 3a, 60-965 Poznań, Polska
agnieszka.weychan@put.poznan.pl

dr inż. Józef Zawodniak
Enea Operator sp. z o.o.
ul. Strzeszyńska 58, 60-479 Poznań, Polska
jozef.zawodniak@enea.pl

Ograniczanie strat energii w żyłach powrotnych linii i mostów kablowych średniego napięcia wykorzystujących kable jednożyłowe

Energy losses' reduction in metallic screens of MV cable power lines and busbar bridges composed of single-core cables

Słowa kluczowe: *linie kablowe, sieć średniego napięcia, straty energii, żyła powrotna*

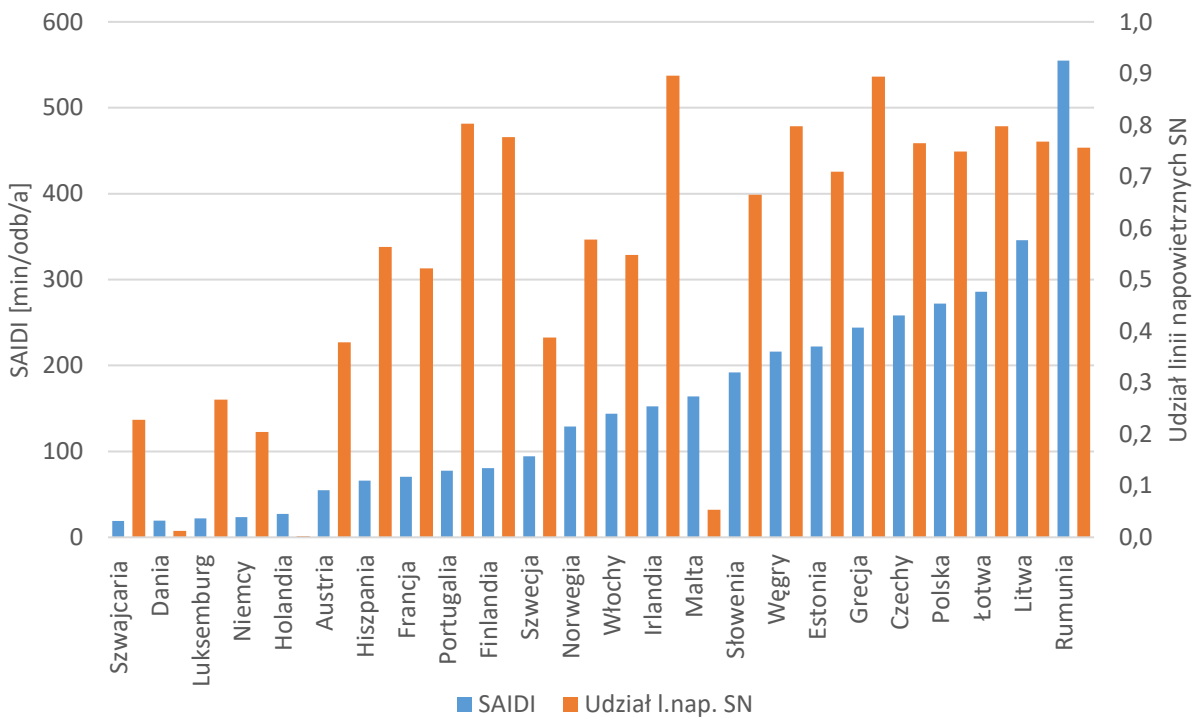
Keywords: *cable lines, medium voltage network, energy losses, metallic cable screen*

Streszczenie: Rosnące skablowanie linii średniego napięcia w sieciach dystrybucyjnych stawia przed operatorami tych sieci wyzwanie prawidłowej eksploatacji linii kablowych. Powiązane jest to z redukowaniem strat energii w żyłach roboczych i powrotnych kabli. W artykule skupiono się na stratach energii w żyłach powrotnych linii oraz mostów kablowych wykonanych przy wykorzystaniu kabli jednożyłowych z metalicznymi żyłami powrotnymi oraz możliwych sposobach ich ograniczania. Przedstawiono analizę symulacyjną i pomiarową poziomu strat energii w żyłach powrotnych kabli wraz z analizą ekonomiczną różnych wariantów ich redukcji poprzez zmianę sposobu pracy tych żył w stosunku do tradycyjnego obustronnego ich uziemienia. Przedstawione zostały również problemy techniczne oraz zagrożenia związane z zastosowaniem rozważanych modyfikacji pracy żył powrotnych podczas zakłóceń zwarciowych w sieciach dystrybucyjnych.

Abstract: The growing share of medium voltage cable lines in distribution networks challenges distribution network operators in terms of proper mode of operation of these lines. It is related to the reduction of energy losses in cable conductors and metallic cable screens. The article focuses on energy losses in metallic cable screens of cable lines and substation busbar bridges composed of single-core cables with metallic screens and possible ways of their reduction. Simulation and measurement analysis of the level of energy losses in the metallic screens of cables is presented together with the economic analysis of various variants of losses reduction through the change of the way these screens are operated in relation to the traditional bilateral earthing at both ends of cable. Technical problems and threats connected with the use of considered modifications of metallic screens operation during earth fault disturbances in distribution networks are also presented.

1. Wstęp

Wskaźniki SAIDI (ang. System Average Interruption Duration Index) charakteryzujące niezawodność dostaw energii elektrycznej z sieci dystrybucyjnej średniego napięcia znacznie odbiegają w Polsce od wskaźników państw europejskich wiodących w tych statystykach. Jedną z istotnych przyczyn tego stanu rzeczy jest stosunkowo niski stopień skablowania odcinków sieci średniego napięcia (SN) nieprzekraczający w Polsce 25% [4]. Spodziewane zyski ze skablowania sieci SN można ocenić porównując średnie wartości wskaźnika SAIDI osiągane w polskich sieciach dystrybucyjnych oraz w sieciach zagranicznych o znacznie większym stopniu skablowania. Sieci dystrybucyjne Szwajcarii, Danii, Luksemburga, Niemiec i Holandii, wykorzystujące zasadniczo sieć kablową, osiągają poziom tego wskaźnika wielokrotnie niższy w porównaniu z sieciami o znacznym udziale linii napowietrznych, co zobrazowano na rys. 1.



Rys. 1. Wskaźniki SAIDI w zestawieniu z udziałem linii napowietrznych w całkowitej długości sieci dystrybucyjnej średniego napięcia na podstawie danych z [4]

W celu poprawy tej sytuacji w Polsce przygotowywany jest szeroki plan skablowania odcinków sieci charakteryzujących się wysokimi statystycznymi wskaźnikami uszkodzeń obejmujący w dużej mierze odcinki przebiegające przez lasy lub tereny zadrzewione. W nachodzącym dwudziestoleciu należy zatem spodziewać się znaczącego przyrostu sieci kablowej średniego napięcia (SN) na poziomie 40 000 km [40]. Przebudowa istniejących linii napowietrznych na linie kablowe pozwoli więc wyeliminować oddziaływanie klimatyczne i zwiększy niezawodność eksploatacyjną tych odcinków [17, 29]. Ponadto, linie kablowe cechują się mniejszym odziaływaniem na krajobraz i są łatwiej akceptowane przez społeczeństwo [49].

W ostatnim dwudziestoleciu nastąpiły też znaczące zmiany w technologii linii kablowych. Najbardziej powszechnym materiałem izolacyjnym stosowanym jest polietylen usieciowany, a do budowy linii kablowych stosuje się kable jednożyłowe z żyłą powrotną mającą zapewnić skuteczne odprowadzenia prądu zwarciowego i umożliwić szybkie działania zabezpieczeń w linii przed długotrwałym oddziaływaniem tego prądu na pozostałe żyły kabla oraz otoczenie [37, 43]. W przypadku braku żyły powrotnej w pobliżu miejsca zwarcia mogłyby wystąpić także istotne zagrożenia porażeniowe. Szczegółowe wymagania w zakresie ochrony porażeniowej w sieciach SN przedstawiono w [14-16].

Przedstawione okoliczności skłaniają do podjęcia badań nad prawidłową eksploatacją sieci dystrybucyjnej średniego napięcia zawierającej coraz większy udział odcinków linii kablowych wykonanych w nowej technologii przy zastosowaniu trzech kabli jednożyłowych z żyłami powrotnymi. Problemem przedstawionym w niniejszym artykule są straty energii w żyłach powrotnych powstające na skutek stosowania ich obustronnego uziemienia oraz sposoby zapobiegania takim stratom na skutek modyfikacji tradycyjnego sposobu pracy żył powrotnych. Przedyskutowano zalety i wady proponowanych modyfikacji. Zagadnienie to przedstawiono wykorzystując komputerowe symulacje pracy linii kablowych, które zweryfikowano na podstawie pomiarów przeprowadzonych w odcinkach rzeczywistych linii kablowych sieci średniego napięcia jednego z polskich operatorów sieci dystrybucyjnych. Zagadnienie wpisuje się w światowe trendy ograniczania strat energii w sieciach dystrybucyjnych [7].

Straty energii w kablach związane są z występowaniem żyły roboczej, izolacji oraz żyły powrotnej lub innych metalowych powłok kabla. Rozróżnić można straty w żyle roboczej związane z jej rezystancją, straty dielektryczne związane z pojemnością kabla i parametrami izolacji oraz straty w żyłach powrotnych związane z przepływem prądu przez te powłoki i prądami wirowymi [2, 35]. Straty związane z prądami wirowymi są zazwyczaj znacznie mniejsze w żyłach powrotnych niż straty związane z przepływem prądu przez obwód żyły powrotnej i nie zależą od układu połączeń w tym obwodzie [37].

Straty energii w liniach kablowych zależą od przekroju znamionowego żyły roboczej i powrotnej, sposobu ułożenia kabli (układ płaski lub trójkąt), a także w dużej mierze od układu połączeń i sposobu uziemienia żył powrotnych kabli [32, 34, 38, 52]. Im większe odległości pomiędzy kablami jednożyłowymi, niezależnie od grubości poszczególnych żył, tym większe obserwuje się straty w tych układach kabli, stąd znacznie większe straty występować będą w przypadku kabli w układzie płaskim, niż w układzie trójkątnym [20, 37]. W przypadku występowania więcej niż jednego systemu kablowego, przykładowo dwóch kabli trójfazowych obok siebie, straty energii zależeć będą również od kolejności faz w poszczególnych kablach jednofazowych [38].

Najbardziej popularne w sieciach SN w Polsce jest obustronne połączenie i uziemienie żył powrotnych na obydwu końcach odcinka linii kablowej. W układach takich, w warunkach

pracy normalnej pod obciążeniem, na żyłach powrotnych występuje pomijalnie niskie napięcie, ale płyną przez nie prądy indukowane przez prądy w żyłach roboczych, przez co występują dodatkowe straty w żyłach powrotnych, a obciążalność prądowa kabli może ulec obniżeniu. Straty w żyłach powrotnych kabli wywołane przez indukowane prądy zależą od impedancji sprzęgających żył roboczych i żył powrotnych kabla [52]. Dla standardowych kabli średniego napięcia dodatkowe straty w żyłach powrotnych stanowią 2-10% całkowitych strat występujących w kablu [20]. Ze względu na dużą liczbę czynników warunkujących straty w żyłach powrotnych, w najbardziej niekorzystnych przypadkach straty w żyłach powrotnych mogą być większe niż straty w żyłach roboczych kabla [18]. Redukcja strat powstających w żyłach powrotnych jest możliwa poprzez stosowanie alternatywnych układów połączeń i uziemienia żył powrotnych, takich jak jednostronne uziemienie żył powrotnych stosowane w kablach wysokiego napięcia [6, 11, 52] i dołączanie ograniczników przepięć [20, 30], przeplecenia żył powrotnych (ang. cross-bonding) [6, 11, 21, 33] lub przeplecenia w połączeniu z transpozycją kabli [44], dołączanie dodatkowych rezystancji lub indukcyjności w miejscu uziemienia żył powrotnych kabli lub w miejscu uziemienia muf przepleceniowych [24, 32, 46] lub zmniejszanie przekroju żył powrotnych, które bardzo często są przewymiarowane [3, 27, 28, 50]. Rozwiązania te rzadko są jednak stosowane w sieciach średniego napięcia ze względu na obawy przed zagrożeniem porażeniowym lub przepięciowym [11, 20] oraz spodziewane problemy eksploatacyjne z wykrywaniem uszkodzeń kabli ze zmodyfikowanymi układami połączeń [10, 23] i możliwość uszkodzeń dodatkowych muf lub szafek kablowych w przypadku przepleceń żył powrotnych [48]. W niniejszym badaniu podjęto nowe zadanie wyznaczania opłacalności stosowania środków ograniczających straty w żyłach powrotnych i badano tę opłacalność w realnych warunkach rynkowych w zależności od obciążenia linii i zastosowanego sposobu przeciwdziałania stratom w żyłach powrotnych.

Proponowane metody redukcji strat w żyłach powrotnych kabli wymagają również analizy podczas stanów zakłóceniowych. Analiza zagrożeń porażeniowych w sieciach kablowych [41] powinna uwzględniać wpływ liczby stacji SN/nn pracujących w sieci SN i ich uziemień oraz rezystancji powłok kabli na rozpływ prądu ziemnozwarciowego. Rozpływ prądu zwarciowego w instalacji uziemiającej jest również przedmiotem analizy w [42], gdzie uwzględnia się wpływ metalowych elementów znajdujących się w ziemi. W [5] analizuje się wpływ parametrów powłok kablowych i uziemień na rozpływ prądu zwarcia jednofazowego i zagrożenia porażeniowe z tym związane. Nowym podejściem zaprezentowanym w niniejszym artykule są badania symulacyjne i poligonowe rozpływu prądu zwarcia jednofazowego w przypadkach odziemiania żył powrotnych w celu identyfikacji stopnia wzrostu zagrożenia porażeniowego w takich przypadkach.

Istotnym elementem analizy pracy kabla w sytuacjach zakłóceniowych jest również badanie przepięć, które mogą występować na odziemionych krańcach żył powrotnych kabla. Zagadnienia dotyczące przepięć są najczęściej analizowane dla linii wysokiego napięcia [6, 20, 48], rzadziej dla linii średniego napięcia [19], jednak ich analiza jest niezwykle istotna dla zapewnienia odpowiedniego stanu izolacji żył powrotnych kabli. W niniejszym artykule zwrócono więc uwagę na możliwe występowanie przepięć przy stosowaniu proponowanych metod redukcji strat w żyłach powrotnych kabli.

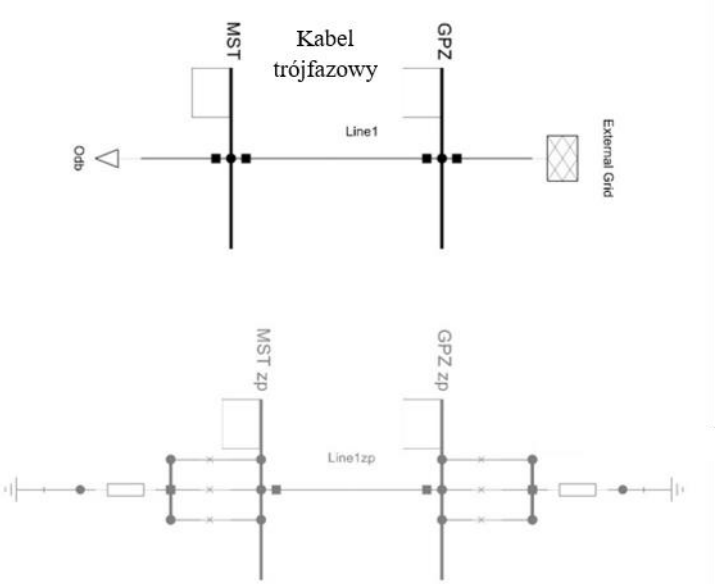
W niniejszym artykule podjęto tę tematykę w celu wypracowania propozycji eksploatacji linii kablowych budowanych w omówionej technologii pozwalającej na redukcję strat w żyłach powrotnych, w uzasadnionych ekonomicznie przypadkach, przy jednoczesnym zachowaniu wymogów w zakresie ochrony przepięciowej, jak również ochrony przeciwporażeniowej na stacjach SN/nn transformujących energię do niższego poziomu

napięcia. W tym celu przeprowadzono szerokie badania symulacyjne sieci średniego napięcia z odcinkami kablowymi, badania na modelach zawierających odcinki rzeczywistych linii kablowych oraz pomiary na odcinkach kabli pracujących w sieci dystrybucyjnej, które opisano poniżej.

2. Modelowanie linii kablowych

Analizowane odcinki linii kablowych pracujących w sieci dystrybucyjnej modelowano przy wykorzystaniu oprogramowania PowerFactory firmy DIgSILENT. W procedurze modelowania linii kablowych tworzony jest najpierw typ kabla jednożyłowego, którego odcinki składają się na kabel trójfazowy. Dane wprowadzane dotyczą wszystkich warstw przewodzących, izolacyjnych i pół-przewodzących, które występują. Wprowadzane są wszystkie parametry geometryczne definiujące przekrój poprzeczny oraz dane definiujące własności wszystkich materiałów składowych.

Posługując się zdefiniowanym typem kabla można utworzyć model symulacyjny systemu kablowego służącego do dystrybucji energii przedstawiony na rys. 2. Danymi uwzględnianymi w tym celu są: typ kabla jednożyłowego wykorzystywanego do budowy systemu kablowego, położenie każdej z żył kabla w stosunku do poziomu ziemi oraz położenie każdej z żył w stosunku do pozostałych żył a także środowisko ułożenia kabla (ziemia, powietrze), częstotliwość, rezystywność gruntu, czy liczba równoległych systemów kablowych.



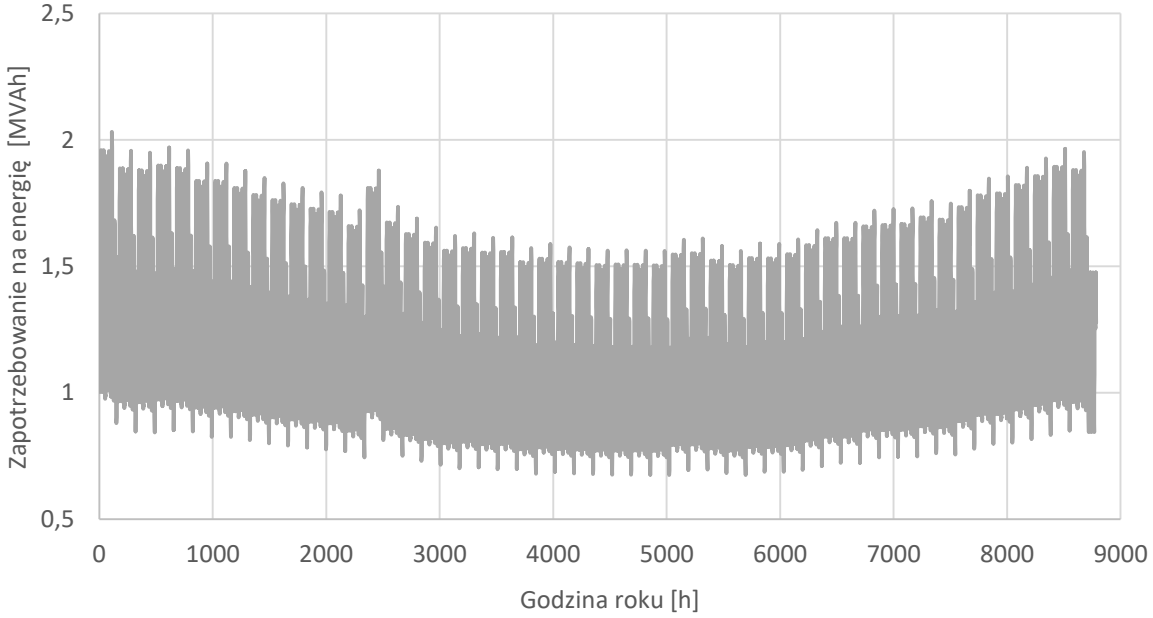
Rys. 2. Model systemu kablowego w sieci

Obliczenia dla systemu kablowego utworzonego przez trzy kable jednożyłowe są prowadzone w oparciu o rozwiazywanie macierzowych równań różniczkowych wiążących prądy w żyłach roboczych i powrotnych kabla z napięciami wzdłuż warstw przewodzących kabla [2, 8, 9]. Pozwala to symulować pracę systemu kablowego w zakresie prądów i napięć w żyłach roboczych i powrotnych w stanach ustalonych i przejściowych podczas normalnej pracy oraz przy zwarciach. Model przedstawiony na rys. 2 umożliwia symulowanie eksploatacji kabla przy różnych sposobach połączenia żył powrotnych ze sobą i z ziemią oraz symulacje

występujących rezystancji uziomów. W przypadku mostów kablowych modelowano system kablowy, jako złożony z trzech lub czterech wiązek kabli w połączeniu równoległym pomiędzy transformatorem WN/SN a szynami zbiorczymi stacji SN. W analizowanym przypadku kable oraz mosty kablowe zamodelowane zostały zgodnie ze standardem kablowych sieci dystrybucyjnych SN przedstawionym w [12].

3. Modelowanie obciążeń i strat energii w ciągach kablowych SN

Zasadniczo celem prowadzonych badań było wyznaczenie strat w żyłach roboczych i powrotnych dla zarejestrowanych obciążeń dla wybranych rzeczywistych odcinków linii kablowych w celu zaproponowania rozwiązań technicznych umożliwiających ograniczanie straty energii w żyłach powrotnych kabla. W sieci miejskiej wyznaczono dwa ciągi kablowe wyprowadzające moc ze stacji WN/SN, określanych jako Główne Punkty Zasilające (GPZ), złożone z trzech odcinków kablowych połączonych szeregowo w układzie trójkątnym. Dla odcinków wychodzących z GPZ zarejestrowano rzeczywiste wartości prądów obciążenia pól liniowych, przekazanych przez Centralną Dyspozycję Mocy (CDM) wykorzystującą centralny system nadzoru nad pracą GPZ-tów, dla wybranych dni 2017 roku w okresie zimowym, wiosennym, letnim oraz jesiennym dla okresów od środy do niedzieli. Na podstawie względnych zmian zużycia energii w poszczególnych tygodniach roku wyznaczony został roczny profil obciążenia dla badanego obiektu. Utworzony roczny profil zapotrzebowania na energię, dla przypadku maksymalnego obciążenia rzędu 2,0 MVA przedstawiony został na rys. 3.

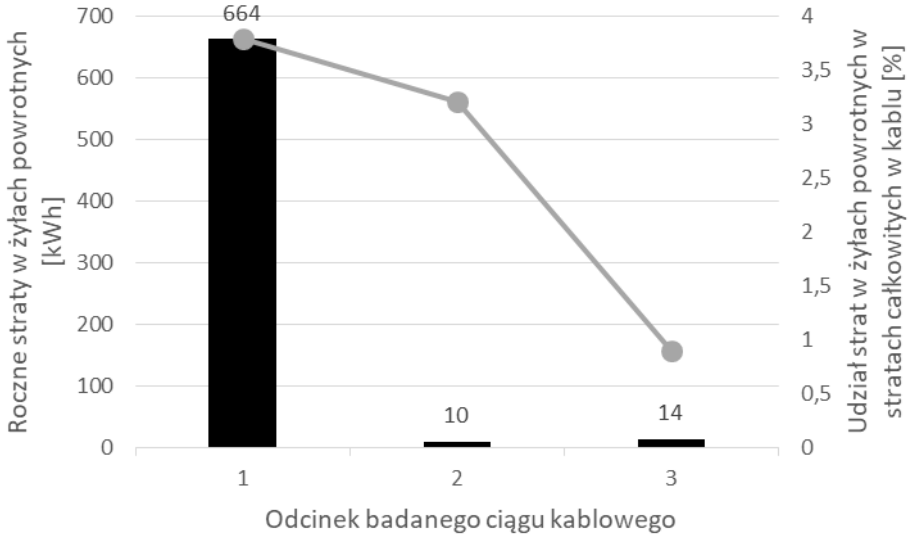


Rys. 3. Roczny profil obciążenia dla linii kablowej wyprowadzającej moc z badanego GPZ-tu

Zobrazowany profil obciążenia dotyczy jedynie odcinka kablowego wyprowadzającego moc z GPZ o przekroju żył roboczych 3x240mm^2 oraz przekroju żył powrotnych 50 mm^2. Dla kolejnych odcinków kablowych przekrój żył roboczych uległ zmniejszeniu do 3x120 mm^2, na skutek zmniejszenia obciążenia ze względu na obecność dodatkowych odpływów w stacjach SN/nn, przy zachowaniu tego samego przekroju żył powrotnych. Uwzględniając liczbę i moce

transformatorów SN/nn na wybranym ciągu zasilania możliwe jest wyznaczenie względnego obciążenia kolejnych odcinków linii kablowych tworzących ciąg przy założeniu równomiernego obciążenia transformatorów rozdzielczych SN/nn, proporcjonalnego do ich mocy znamionowych. Wyznaczono w ten sposób obciążenie drugiego odcinka ciągu na poziomie 72% obciążenia pierwszego odcinka. Analogicznie, uwzględniając kolejne miejskie stacje zasilające odcinek trzeci był obciążony na poziomie 63% obciążenia pierwszego odcinka. W miarę rozwoju systemów inteligentnego opomiarowania dokładniejszą metodą wyznaczenia obciążeń poszczególnych odcinków linii kablowej byłoby wykorzystanie danych dotyczących obciążeń stacji SN/nn na podstawie danych z liczników bilansujących zainstalowanych na tych stacjach, lecz dane takie nie były dostępne w czasie prowadzenia przedstawianych badań.

Dla pierwszego z rozważanych ciągów, z uwagi na duże różnice w długości odcinków kabla, straty w odcinku II oraz III były znacznie mniejsze niż w odcinku I. Straty w poszczególnych odcinkach badanego ciągu kablowego wykazują się zmiennością sezonową, zgodnie ze zmiennością obciążenia zamodelowaną na rys. 3. Wykonane analizy pozwalały na określanie strat w żyłach powrotnych kabla, przedstawionych w formie wykresu na rys. 4. jako roczne bezwzględne wartości strat w kWh oraz jako udział w całkowitych stratach energii w kablu dla analizowanych odcinków. Z analizy wynika, że średni udział strat energii w żyłach powrotnych wynosił 3,5% strat całkowitych w kablu.

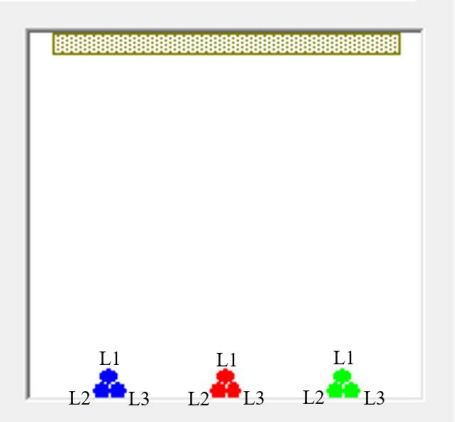


Rys. 4. Poziom strat energii w żyłach powrotnych badanych odcinków linii kablowej

Dla drugiego badanego ciągu kablowego przeprowadzono analogiczne symulacje. Moc roczna maksymalna wyprowadzana tym ciągiem z GPZ, przy tym samym przekroju linii kablowej jak dla pierwszego ciągu, była znacząco niższa rzędu 0,8 MV, lecz drugi odcinek ciągu był obciążony tak samo jak pierwszy natomiast trzeci w stopniu nieznacznie niższym rzędu 88% obciążenia pierwszego odcinka. Ze względu na mniejsze obciążenie drugiego badanego ciągu kablowego wartości bezwzględne strat w żyłach powrotnych kabla są mniejsze i dla całego odcinka wynoszą ok. 40 kWh rocznie. Udział strat w żyłach powrotnych w całkowitych stratach w poszczególnych odcinkach kabli wynosi 2,3% dla pierwszego odcinka, 1,1% dla drugiego oraz 0,5% dla trzeciego, wynosząc średnio 1,5%.

Przeanalizowano również pod względem strat w żyłach powrotnych straty w moście kablowym zbudowanym z trzech odcinków kabli trójfazowych 3x3x240/50 mm^2 połączonych

równolegle, o długości 23 m. Schemat ułożenia poszczególnych kabli jednożyłowych w badanym moście kablowym przedstawiono na rys. 5. Warto zauważyć, że podobne układy pracy linii kablowych można spotkać w liniach wyprowadzających moc z dużych lokalnych źródeł energii [19]. Obciążenie maksymalne mostu wyprowadzającego moc z transformatora 25 MVA wynosiło 13 MVA. Po uwzględnieniu sezonowych zmian obciążenia roczne straty w żyłach powrotnych na podstawie symulacji obliczeniowych określono na poziomie 150 kWh, co stanowi ok. 6,7% całkowitych strat w rozpatrywanym moście kablowym.



Rys. 5. Schemat budowy mostu kablowego

4. Weryfikacja pomiarowa wyników symulacji

Weryfikacja pomiarowa modeli obliczeniowych linii kablowych SN wymagała przeprowadzania serii pomiarów dla przykładowych odcinków linii kablowych oraz mostów kablowych zbudowanych z trzech kabli jednożyłowych. Celem weryfikacji było stwierdzenie poziomu rozbieżności w wartościach prądów w żyłach powrotnych wynikających z pomiarów prądów w żyłach roboczych, w stosunku do wyników symulacyjnych otrzymywanych poprzez modelowanie prądów w żyłach powrotnych przy wykorzystaniu oprogramowania firmy DIgSILENT.

Pomiary weryfikacyjne zostały wykonane na trzech odcinkach linii kablowych ułożonych w układzie trójkątnym oraz na moście kablowym na stacjach z łatwą dostępnością żył powrotnych. Dla każdego z powyższych odcinków zmierzone zostały wartości prądów w żyłach roboczych kabla oraz prądów w żyłach powrotnych. Do wykonania pomiarów wykorzystane zostały następujące urządzenia pomiarowe:

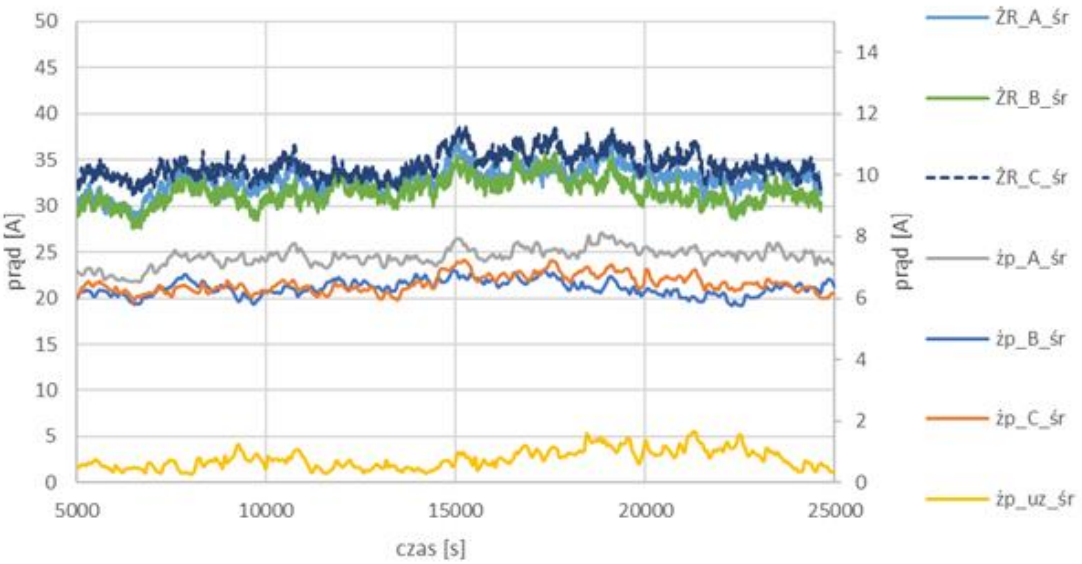
- analizator jakości zasilania i energii Fluke 435 dla żył roboczych badanych odcinków kablowych wyposażony w cewki Rogowskiego – dokładność pomiaru 2%,
- mierniki cęgowe BRYMEN BM135s dla żył powrotnych badanych odcinków kablowych – dokładność pomiaru 5%,
- w przypadku mostów kablowych, jako wartość prądu w żyłach roboczych przyjęto wartości prądów rejestrowane w systemie nadzoru nad pracą GPZ, podawane z dokładnością do czwartej cyfry znaczącej.

Błąd określano zgodnie z poniższą zależnością:

$$\delta I_{żp} = \frac{I_{pom\_żp} - I_{sym\_żp}}{I_{pom\_żp}} \cdot 100\% \tag{1}$$

gdzie: $I_{pom\_żp}$ – wartość prądu w żyle powrotnej kabla wyznaczona z pomiarów (średnia wartość w ciągu minuty), $I_{sym\_żp}$ – wartość prądu w żyle powrotnej kabla dla analogicznego stanu obciążenia uzyskana w programie PowerFactory.

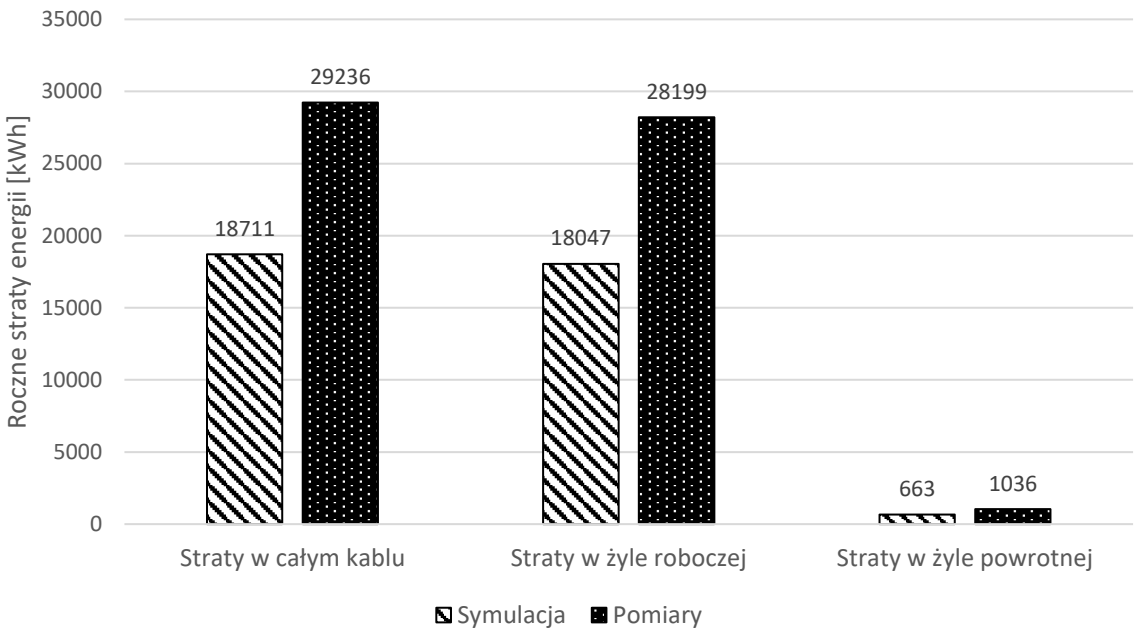
Zarejestrowane pomiary prądów w żyłach roboczych i powrotnych przeanalizowano dla poszczególnych stanów obciążeń, określonych z częstością uśredniania wartości skutecznych zastosowanych mierników. Przykładowe wyniki rejestracji przedstawiono na rys. 6.



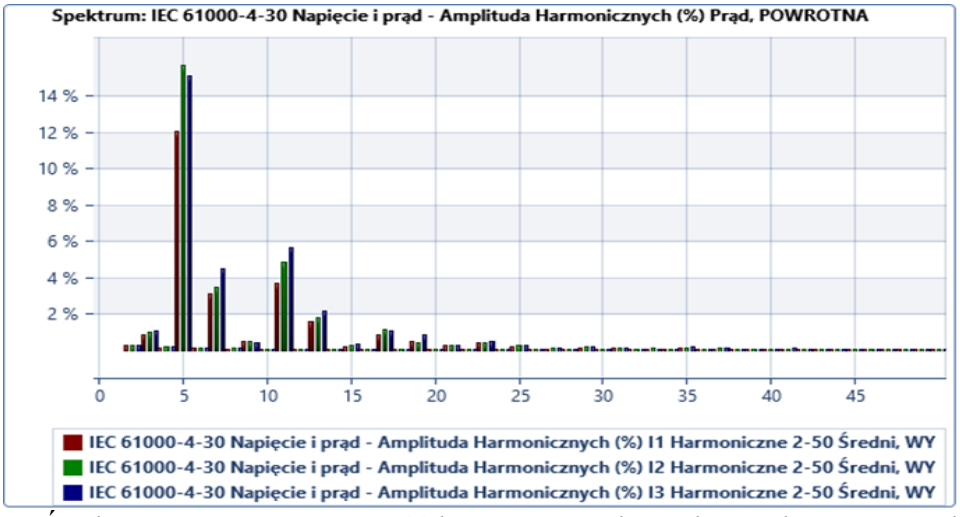
Rys. 6. Prądy w żyłach roboczych i powrotnych w drugim okresie pomiarów (ŻR- żyła robocza; żp - żyła powrotna; uz – uziom żył powrotnych)

Stwierdzono znaczącą rozbieżność wyników symulacji oraz wartości prądów pomierzonych dla analizowanych obiektów wynoszącą dla prądów rzeczywiście mierzonych w żyłach powrotnych kabli od +20% do +35%. Biorąc pod uwagę weryfikację pomiarową symulacji komputerowych prądów w żyłach powrotnych oraz błąd pomiaru zastosowanych przyrządów pomiarowych w dalszych rozważaniach dotyczących opłacalności stosowania środków zaradczych dla ograniczania poziomu strat w żyłach powrotnych założono, że prądy płynące w żyłach powrotnych kabli są o 25% większe w stosunku do wyników otrzymanych w symulacjach komputerowych. Na rysunku 7 przedstawiono wzrost strat rocznych, w wyniku przyjętego na podstawie pomiarów odchylenia rzeczywistego prądu w żyłach powrotnych w stosunku do symulowanego (dla odcinka linii kablowej analizowanego na rys. 3).

Jedną z przyczyn powstałych błędów może być zawartość harmonicznych w prądach żył roboczych, przenosząca się na żyły powrotne przedstawiona na rys. 8, a nieuwzględniana w obliczeniach symulacyjnych. Kolejną przyczyną mogą być prądy płynące przez żyły powrotne, będące wynikiem różnic potencjałów stacji łączonych przez odcinek badany. Prąd płynący przez uziom, przedstawiony na rys. 6, jest sumą tych prądów i obejmuje prądy harmoniczne spływające do uziomu.



Rys. 7. Porównanie rocznych wartości strat w żyłach powrotnych badanych odcinków linii kablowych otrzymanych z symulacji oraz po weryfikacji pomiarami



Rys. 8. Średnia miesięczna zawartość harmonicznych prądu w żyle powrotnej kabla wyprowadzającego moc z jednego z GPZ

5. Dobór przekroju i modyfikacje sposobu uziemiania żył powrotnych kabla prowadzące do ograniczania strat mocy czynnej

Występujące straty w żyłach powrotnych odcinków linii kablowych można ograniczyć następującymi sposobami:

- przeplecenia,
- odziemienia jednostronne,
- zastosowanie niższego przekroju w stosunku do przyjmowanego standardowo.

Pierwszy z wymienionych sposobów, w przypadku zastosowania dwóch przepleceń na jednej trzeciej i dwóch trzecich długości kabla, zapewnia ograniczenie strat praktycznie do wartości pomijalnych przy prądach nieodkształconych. Wymaga on jednak zastosowania

specjalnych muf przepleceniowych lub budowy skrzynek kablowych dla dokonywania przepleceń, co jest kosztowne. Służby eksploatacyjne kabli obawiają się także przepleceń ze względu na oczekiwane utrudnienia w prowadzeniu eksploatacyjnej diagnostyki linii kablowych.

Odziemianie jednostronne żył powrotnych ogranicza straty, lecz wymaga analizy i stwierdzenia dopuszczalności występujących wartości przepięć izolacji żyły powrotnej na odziemionych końcach kabla w przypadku prądów zwarciowych płynących przez żyły robocze. Ponadto uziemienie żył powrotnych stosowane jest dla ograniczenia prądu uziomowego spływającego przez uziom roboczy stacji SN/nn zasilanych liniami kablowymi, co przyczynia się do poprawy bezpieczeństwa porażeniowego na stacjach SN. Zatem, w przypadku linii kablowych, konieczna jest także weryfikacja wartości napięć rażeniowych występujących przy zwarciach jednofazowych na tych obiektach w przypadku odziemienia jednej czy dwóch żył powrotnych.

Mniejszy przekrój żył powrotnych, to rozwiązanie techniczne do zastosowania na etapie budowy nowych odcinków linii kablowych, które prowadzi do znacznych oszczędności w nakładach inwestycyjnych. Ponadto wyższa wartość rezystancji żyły powrotnej powoduje ograniczenie prądu indukowanego w żyłach dwustronnie uziemionych, co prowadzi do ograniczenia strat. Zagrożeniem dla żył o zmniejszonym przekroju może być niedopuszczalny wzrost temperatury przy zwarciach dwufazowych z udziałem ziemi na odcinku wyprowadzającym moc z GPZ [26]. Można jednak w takim przypadku, dla zapobieżenia ich szkodliwemu oddziaływaniu cieplnemu na żyły powrotne, uruchomić na takim odcinku bezzwłoczne zabezpieczenia nadprądowe.

Zastosowanie jednego z wyżej wymienionych sposobów ograniczenia strat energii w żyłach powrotnych kabli powinno być efektywne ekonomicznie. Stanem wyjściowym pozwalającym ocenić uzyskane oszczędności są straty występujące w obecnie stosowanym standardzie pracy kabla przy obustronnie uziemionych żyłach powrotnych. Oszczędności uzyskane dzięki zastosowanej modernizacji powinny w okresie eksploatacji kabla przewyższać nakłady inwestycyjne pozwalające na wdrożenie określonego środka zaradczego służącego ograniczaniu poziomu strat, co można wyrazić poniższą zależnością:

$$\Delta E_{Lmd} - V_i > 0 \tag{2}$$

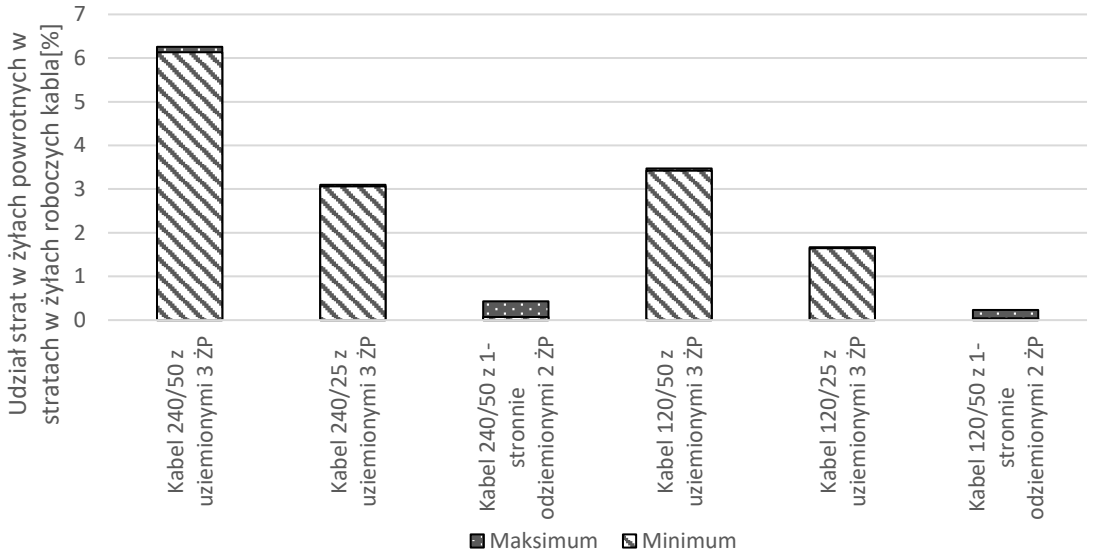
gdzie: ΔE_{Lmd} – zdyskontowane, za rozważany czas eksploatacji, oszczędności na skutek modyfikacji sposobu pracy żył powrotnych kabla, V_i – nakłady na modernizację pracy żył powrotnych kabla.

Podejmowanie decyzji o zastosowaniu konkretnego sposobu ograniczania strat w żyłach powrotnych powinno być uzasadnione uproszczonymi analizami ekonomicznymi wykorzystującymi stosunkowo łatwo dostępne dane dotyczące analizowanej linii kablowej. W celu oceny możliwych do osiągnięcia oszczędności na stratach w żyłach powrotnych kabla przeprowadzono symulacje obliczeniowe strat w żyłach powrotnych w zależności od stopnia obciążenia kabli w przypadku zastosowania przepleceń, odziemiania jednostronnego tych żył oraz zastosowania zmniejszonego przekroju. Uwzględniono wyniki weryfikacji pomiarowej wykazującej 25% zwiększenie prądu w żyłach powrotnych w stosunku do wartości uzyskanych w symulacjach obliczeniowych. Przekłada się to na następujący wzrost strat w żyłach powrotnych (3):

$$\Delta P_{rz} = I_{rz}^2 \cdot R = (1{,}25 \cdot I_{sym})^2 \cdot R = 1{,}5625 \cdot I_{sym}^2 \cdot R = 1{,}5625 \cdot \Delta P_{sym} \tag{3}$$

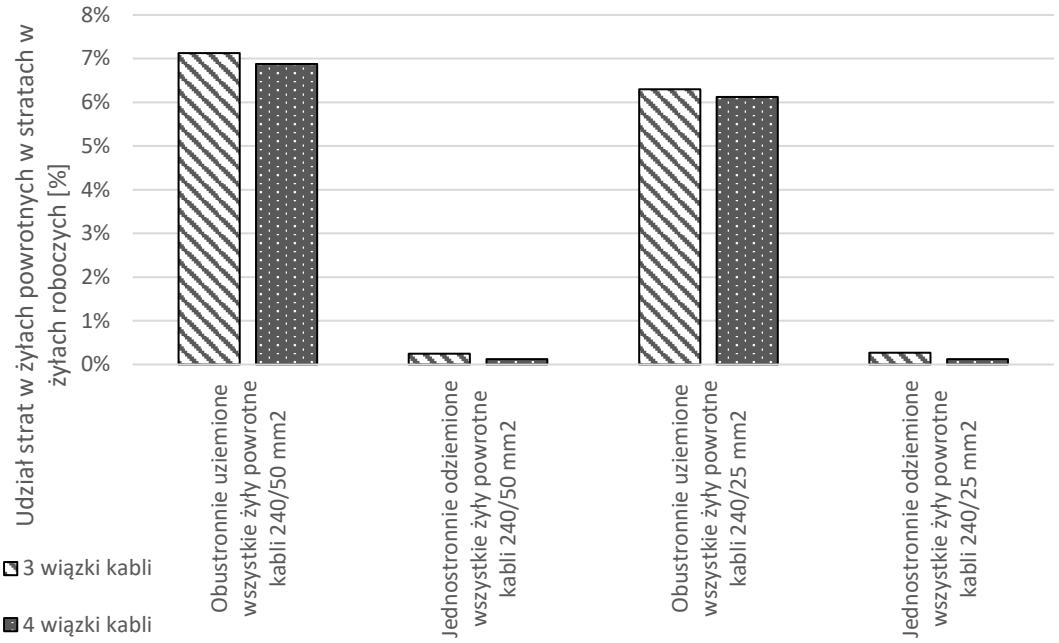
gdzie: ΔP_{sym} – straty mocy otrzymane z symulacji, I_{sym} – prąd w żyłach powrotnych otrzymany z symulacji, R – rezystancja kabla, I_{rz} – prąd rzeczywisty w żyłach powrotnych kabli, ΔP_{rz} – rzeczywiste straty w żyłach powrotnych kabli.

W przypadku zastosowania przepleceń straty w żyłach powrotnych są pomijalnie małe. Dla pozostałych analizowanych przypadków wyniki obrazujące udziały strat w żyłach powrotnych w stosunku do strat w żyłach roboczych kabla dla wybranych przekrojów kabli i żył powrotnych przedstawiono na rys. 9. Widoczny zacieniony zakres wyników dla poszczególnych rodzajów kabli wynika z różnych długości odcinków kablowych analizowanych leżących w zakresie 0,1 do 1 km. Dla przeprowadzonych symulacji stwierdzono, że udział strat w żyłach powrotnych w stratach w żyłach roboczych kabla jest stały dla poszczególnych układów kabli w zakresie od 5 do 100% obciążenia prądem znamionowym.



Rys. 9. Udział strat mocy w żyłach powrotnych w stratach w żyłach roboczych dla kabli 240 oraz 120 mm^2 dla przypadków odziemiania i ograniczenia przekroju tych żył

W przypadku krótkich odcinków linii kablowych pracujących w mostach kablowych przeplecenia nie mają zastosowania, natomiast racjonalne jest odziemienie jednostronne wszystkich żył powrotnych, jako że uziemienie żył powrotnych nie ma istotnego znaczenia ze względu na bezpieczeństwo porażeniowe na stacji WN/SN po poprawnym zaizolowaniu odziemionych końców. Udział strat mocy w żyłach powrotnych w stratach w żyłach roboczych mostów kablowych dla różnych konfiguracji połączeń żył powrotnych oraz analizowanych mostów kablowych uzyskane na podstawie skorygowanych pomiarowo symulacji pracy mostów przy różnych obciążeniach przedstawiono na rys. 10.



Rys. 10. Udział strat mocy w żyłach powrotnych w stratach w żyłach roboczych w mostach kablowych dla różnych konfiguracji połączeń oraz mostów kablowych złożonych z 3 lub 4 wiązek kabli

Przedstawione wyniki można wykorzystać dla oszacowania oszczędności rocznych wynikających ze zmiany sposobu uziemiania żył powrotnych zgodnie z niżej przedstawioną procedurą.

a. Wyznaczanie strat maksymalnych P_{max} w kablu w żyłach roboczych (4):

$$P_{max} = 3 \cdot I_{max}^2 \cdot \frac{l}{\gamma \cdot s} \tag{4}$$

gdzie: I_{max} – prąd w żyle roboczej odcinka kabla analizowanego, l – długość odcinka [m], γ – przewodność żyły roboczej [m/(Ωmm^2)], s - przekrój [mm^2].

Dane dotyczące obciążenia analizowanych odcinków można pozyskać z rejestracji obciążeń pól w GPZ przez system zarzadzania pracą sieci (DMS) dla wyprowadzeń kablowych z GPZ; dla ciągów w głębi sieci można dokonać podziału prądu kabla wyprowadzającego moc z GPZ, zarejestrowanego przez DMS, wg udziału sumy mocy transformatorów SN/nn zasilanych z danego kabla do sumy mocy wszystkich transformatorów zasilanych z analizowanego ciągu wyprowadzającego moc z GPZ.

b. Określenie rocznych oszczędności w stratach energii $\Delta E_{l(BEB\text{-}m)i}$ w żyłach powrotnych w roku i na skutek zmiany sposobu ich pracy z obustronnego uziemienia (ang. both-end earthing *BEB*) ma analizowany sposób po modernizacji (m):

$$\Delta E_{l(BEB-m)i} = (\Delta P_{lBEB\%} - \Delta P_{lm\%}) \cdot P_{max} \cdot \tau_{max} \cdot \left(1 + \frac{\Delta P_a}{100}\right)^i \tag{5}$$

gdzie: $\Delta P_{lBEB\%}$; $\Delta P_{lm\%}$ - wartości strat mocy [%] dla rozważanych sposobów pracy żył powrotnych przedstawione przykładowo na rys. 9 oraz 10; P_{max} - straty w żyłach roboczych odcinka linii kablowej według zależności (4), τ_{max} – roczny czas trwania strat maksymalnych w analizowanym odcinku kabla, ΔP_a - przyrost roczny strat na skutek przyrostu obciążenia odcinka [%].

W celu określenia rozpływów mocy w poszczególnych odcinkach linii kablowych P_{max} można posłużyć się metodą rozdziału obciążeń na podstawie mocy znamionowych zasilanych transformatorów, rozdziału obciążeń na podstawie miesięcznych odczytów przepływającej energii, analizą obciążeń na podstawie standardowych profili obciążenia lub wykorzystaniu danych rejestrowanych przez liczniki inteligentne (ang. AMI, Advanced Metering Infrastructure) [1]. Powszechnie uważa się, że wykorzystanie danych z systemu AMI pozwala na osiągnięcie najdokładniejszych wyników [25, 47], niemniej jednak ze względu na fakt, że system AMI nie osiągnął pełnej funkcjonalności w analizach wykorzystuje się często dane wynikające ze standardowych profili obciążeń [22].

Czasy trwania strat maksymalnych τ_{max} dla odcinków sieci dystrybucyjnej w Polsce podawane są w literaturze i zgodnie z [36] dla mostów kablowych czy odcinków wyprowadzających moc z pól GPZ zakres tych czasów leży w przedziale 1248 – 4449 h przy wartości średniej 2525 h i odchyleniu standardowym 510 h. Wybór właściwego czasu trwania strat maksymalnych zależy od rodzajów obciążeń przenoszonych danym elementem – większe czasy dla obciążeń przemysłowych, mniejsze dla bytowo-komunalnych. Dla odcinków w głębi sieci, obciążonych stacjami SN/nn, zakres rozważanych czasów leży w obszarze wartości od 788 do 2444 h przy wartości średniej 1662 h i odchyleniu standardowym 801 h. Dla ciągów liniowych zasilających poszczególne stacje SN/nn w głębi sieci, gdy nie dysponujemy obciążeniami z pomiarów, racjonalne będzie przyjęcie wartości z podanego zakresu z uwzględnieniem udziału sumy mocy znamionowych zasilanych z danej sekcji kabla do sumy mocy wszystkich stacji transformatorowych zasilanych z analizowanego ciągu liniowego, a także rodzaju zasilanych odbiorów.

c. Zdyskontowana wartość strat za okres eksploatacji kabla ΔE_{lmd}

W celu wyznaczenia wartości zaoszczędzonych strat zdyskontowanych ΔE_{lmd} należy zsumować oszczędności roczne powstałe w wyniku modernizacji, których wartość sprowadzamy do roku zerowego, stosując poniższą zależność:

$$\Delta E_{Lmd} = \sum_{i=1}^{25} \left[\Delta E_{l(BEB-m)i} \cdot C_{ee0} \cdot \left(1+\frac{\Delta C_a}{100}\right)^{i} \cdot \left(1+\frac{R_d}{100}\right)^{-i} \right] \tag{6}$$

gdzie: $\Delta E_{l(BEB-m)i}$ - roczne oszczędności w stratach energii dane zależnością (5), C_{ee0}, ΔC_a – cena rynkowa energii elektrycznej w roku zerowym [zł/kWh] i jej spodziewany roczny przyrost w [%], R_d – roczna stopa dyskonta w [%] dla określenia wartości strat w roku 0.

Okres dyskonta przyjęto na 25 lat to znaczy nieznacznie powyżej księgowego okresu amortyzacji linii kablowej wynoszącego 22,5 roku [45].

Nakłady inwestycyjne na modernizację sposobu pracy żył powrotnych odcinków kabli V_i występujące w nierówności (2) zależą od zastosowanego rozwiązania. Najtańszym rozwiązaniem jest jednostronne uziemienie żył powrotnych, co wiąże się z kosztami pracy brygady eksploatacyjnej przeprowadzającej zaizolowanie odziemionego końca kabla i niewielkimi kosztami materiałowymi. Zastosowanie przepleceń jest zdecydowanie bardziej kosztowne, ponieważ wymaga zakupienia i zainstalowania muf przepleceniowych lub skrzynek kablowych dla realizacji przepleceń. Zastosowanie żyły powrotnej o mniejszym przekroju, w stosunku do dotychczas uznawanego za rozwiązanie standardowe, prowadzi do znaczących oszczędności w nakładach inwestycyjnych, które są tym większe im dłuższy jest analizowany odcinek kabla oraz skutkuje ograniczeniem poziomu strat, ale jest do wdrożenia tylko na etapie budowy linii kablowej, bo jej wymiana na linię o mniejszym przekroju żyły powrotnej nie jest

ekonomicznie efektywna z powodu wysokich nakładów inwestycyjnych zbliżonych do rozwiązania pierwotnego budowy linii kablowej.

Przeprowadzono wiele analiz opłacalności zastosowania przedstawionych wyżej modernizacji sposobu pracy żył powrotnych dla ograniczenia strat w żyłach powrotnych kabli. Koszty strat sprowadzano do roku zerowego, tzn. 2017 przy zastosowaniu dwóch stóp dyskonta: niskiej, wynoszącej 2,5% stosowanej przy obliczaniu wartości pomocy publicznej w przypadku rozłożenia płatności na raty oraz wyższej, wynikającej z kosztu kapitału operatorów sieci dystrybucyjnych w Polsce wynoszącej 5,633%. Założono wzrost rocznego obciążenia odcinków na poziomie 0,5% i roczny wzrost cen energii na pokrycie strat na poziomie 2,5% w stosunku do ceny z 2017. Wyniki przeprowadzonych analiz skłaniają do sformułowania następujących wniosków:

- brak opłacalności stosowania przepleceń istniejących odcinków linii kablowych ze względu na niską wartość zdyskontowanych strat zaoszczędzonych, wynoszącą dla analizowanych odcinków linii maksymalnie do 3000 zł w okresie 25 lat, dla bieżących obciążeń linii kablowych charakterystycznych dla odbiorców bytowo-komunalnych,
- opłacalność zastosowania odziemiania dwóch żył powrotnych dla kabli obciążonych mocą maksymalną rzędu 2 MVA dla odbiorców bytowo-komunalnych,
- brak opłacalności zastosowania odziemiania żył powrotnych w istniejących mostach kablowych ze względu na niską wartość zdyskontowanych strat w żyłach powrotnych tych mostów obciążenia, co wynika z ich małej długości, mimo ich znaczącego obciążenia,
- opłacalność zastosowania w nowych liniach kablowych żył powrotnych o przekroju 25 mm^2 w miejsce standardowo stosowanych dotychczas 50 mm^2 i rozważanie opłacalności odziemienia 2 z 3 żył powrotnych takich kabli w zależności od wyników szacunkowej analizy ekonomicznej,
- opłacalność zastosowania w mostach kablowych żył powrotnych o przekroju 25 mm^2 przy jednostronnym odziemieniu wszystkich żył powrotnych w przypadku nowych inwestycji.

6. Zagrożenia wynikające z odziemiania żył powrotnych odcinków linii kablowych

Jednostronne odziemianie wybranych lub wszystkich żył powrotnych kabli jest najbardziej interesującym sposobem zapobiegania nadmiernym stratom w żyłach powrotnych kabli. Niestety zidentyfikowano również zagrożenia wynikające z tego sposobu pracy żył powrotnych omówione poniżej.

6.1. Zagrożenia porażeniowe na stacjach rozdzielczych SN/nn

Odziemianie żył powrotnych wpływa na prądy uziomowe oraz na napięcia rażeniowe mogące występować podczas zwarć doziemnych na stacjach SN/nn oraz w obwodach nn zasilanych z tych stacji [41, 42]. Prąd zwarciowy doziemny I_{k1} rozdziela się na składową spływającą poprzez uziom stacji I_e oraz prąd powracający żyłami powrotnymi I_s:

$$I_{k1} = I_e + I_s \tag{7}$$

Tylko składowa I_e powoduje wzrost napięć rażeniowych na stacjach. Zmniejszanie prądu uziomowego I_e, na skutek odprowadzania części prądu zwarciowego przez żyły powrotne kabli, wyznacza się w oparciu o znajomość wartości współczynnika redukcyjnego r powodującego ograniczenie wartości prądu uziomowego, zgodnie ze wzorem:

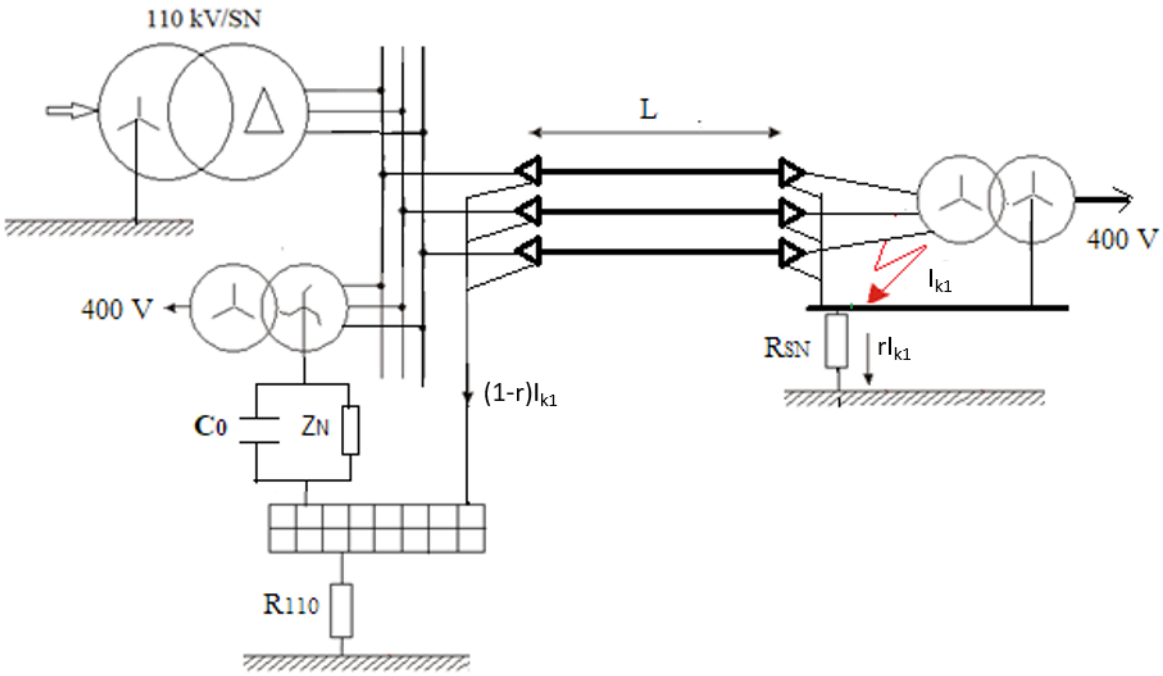
$$I_e = r \cdot I_{k1} \tag{8}$$

Współczynnik redukcyjny wyznaczany dla linii kablowej złożonej z kabli jednożyłowych jest najczęściej opisywany zależnością:

$$r = 1 - \frac{Z_M}{Z_{ŻP}} = \frac{Z_{ŻP} - Z_M}{Z_{ŻP}} \tag{9}$$

gdzie: Z_M – oznacza impedancję wzajemną pomiędzy żyłą roboczą, w której przepływa prąd I_k a żyłami powrotnymi kabla, $Z_{ŻP}$ – oznacza wypadkową impedancję własną żył powrotnych kabla.

Wzór ten jest w pełni uzasadniony przy pominięciu rezystancji uziemienia trzech żył powrotnych. Po uwzględnieniu rezystancji uziemień [5], rozpływ prądu zwarcia doziemnego będzie zgodny z rys. 11, na którym zaznaczono przypadek zakłócenia doziemnego w stacji SN/nn zasilanej linią kablową bezpośrednio z GPZ-tu.



Rys. 11. Rozpływ prądu zwarcia doziemnego na stacji SN/nn z uwzględnieniem żył powrotnych kabla zasilającego oraz rezystancji uziemień obu stacji; Z_N – impedancja uziemienia punktu neutralnego sieci SN (rezystor lub cewka Petersena), C_0 - pojemność doziemna sieci, I_{k1} – prąd zwarcia doziemnego, R_{110} – rezystancja uziemienia głównej stacji zasilającej, R_{SN} – rezystancja uziemienia stacji odbiorczej, L – długość rozpatrywanego odcinka kabla [km]

Różnicę potencjałów pomiędzy uziomami GPZ i stacji SN/nn opisać można następującą zależnością:

$$(1-r)I_{k1}Z_{ŻPj} - I_{k1}Z_{Mj} \cdot L = rI_{k1}(R_{110} + R_{SN}) \tag{10}$$

gdzie: Z_{Mj} i $Z_{ŻPj}$ - impedancje jednostkowe [Ω/km], r – współczynnik redukcyjny, R_{110} – rezystancja uziemienia głównej stacji zasilającej, R_{SN} – rezystancja uziemienia stacji odbiorczej, L – długość rozpatrywanego odcinka [km].

Przekształcając wzór (10) otrzymuje się zależność na prąd uziomowy w postaci:

$$I_e = I_{k1} \cdot \frac{(Z_{ŻPj} - Z_{Mj}) \cdot L}{(R_{110} + R_{SN}) + Z_{ŻPj} \cdot L} \tag{11}$$

Ze względu na fakt, że najczęściej $R_{110} << R_{SN}$ w obliczeniach uwzględnia się przede wszystkim rezystancję uziemienia stacji odbiorczej R_{SN}.

Jak łatwo zauważyć powyższą zależność można również przedstawić w postaci:

$$I_e = r \cdot I_{k1} \cdot \frac{Z_{ŻPj} \cdot L}{R_{SN} + Z_{ŻPj} \cdot L} \tag{12}$$

lub

$$I_e = r_{rz} \cdot I_{k1} \tag{13}$$

gdzie:

$$r_{rz} = r \cdot K_{KOR} \tag{14}$$

a współczynnik korekcyjny K_{KOR}:

$$K_{KOR} = \frac{Z_{ŻPj} \cdot L}{R_{SN} + Z_{ŻPj} \cdot L} \tag{15}$$

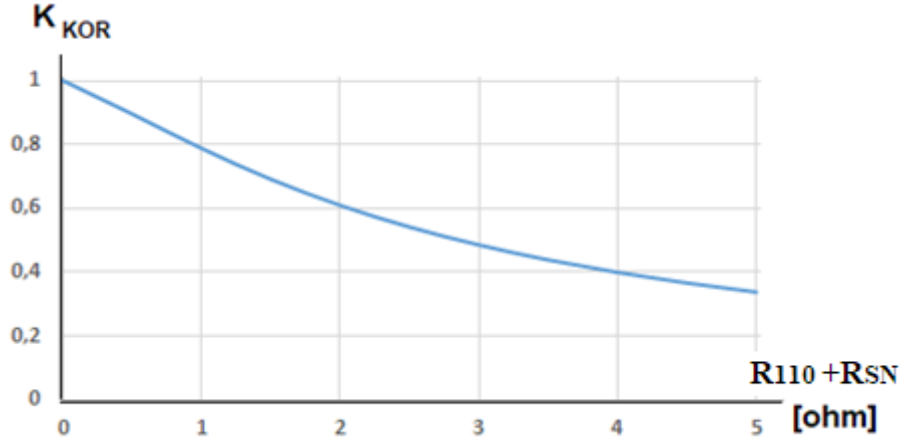
Współczynnik korekty K_{KOR} we wzorze (15) dotyczy układu, w którym uziemione są trzy żyły robocze. Po odziemieniu jednej lub dwóch żył powrotnych w stacji SN/nn wartości tego współczynnika ulegną zmianie, ponieważ zmienią się wartości impedancji $Z_{ŻP}$. Zakres spodziewanych zmian przedstawiono w tabeli 1 dla trzech sposobów wyznaczania współczynnika korekcyjnego:

- z pomiarów na odcinku linii kablowej zasilonej napięciem obniżonym 230/400 V,
- uproszczonych obliczeń analitycznych uwzgledniających układ przestrzenny odcinka linii kablowej i jej parametry,
- symulacji obliczeniowych przy zastosowaniu oprogramowania PowerFactory.

Tabela 1. Zmiany współczynnika redukcyjnego przy odziemianiu żył powrotnych w stosunku do jego wartości podczas uziemienia trzech żył przy założeniu porównywalnych wartości impedancji $Z_{ŻP}$ i rezystancji uziemień (R_{110}+R_{SN})

Lp.	Wskaźnik wzrostu wartości *r* wg pomiarów	Wskaźnik wzrostu wartości *r* wg uproszczonych obliczeń	Wskaźnik wzrostu wartości *r* wg symulacji PF	Uwagi
1	1,20	1,28	1,20	po odziemieniu w stacji SN/nn jednej żyły powrotnej
2	1,70	1,79	1,75	po odziemieniu w stacji SN/nn dwóch żył powrotnych

Przy ocenie rozpływu prądów uziomowych w stacjach SN zasilanych liniami kablowymi należy w obliczeniach współczynnika redukcyjnego uwzględniać również wpływ rezystancji uziemień stacji. Dla przyjętej sumarycznej rezystancji uziemień stacji ($R_{110}+R_{SN}$) w zakresie 0-5 Ω korekta współczynnika redukcyjnego zmieniać się może zgodnie z krzywą przedstawioną na rys. 12. W celu określenia rezystancji uziomów rzeczywistej stacji SN/nn należy wykonać pomiary zgodnie z [13, 14].



Rys. 12. Zmiany współczynnika korygującego wartości współczynnika redukcyjnego w zależności od sumarycznej rezystancji uziemień żył powrotnych kabla

Badania symulacyjne i próby poligonowe na linii kablowej przy napięciu obniżonym wykazały zwiększanie współczynnika redukcyjnego w następstwie odziemiania żył powrotnych. Dla analizowanego przypadku, wartości tego współczynnika dla obustronnie uziemionych żył powrotnych są na poziomie 0,5 i wzrastają do wartości 0,6 po odziemieniu jednej żyły powrotnej lub do 0,85 po odziemieniu dwóch żył. Dopuszczalność stosowania odziemienia żył powrotnych powinna być zatem weryfikowana pod względem dopuszczalnych napięć rażeniowych na stacji, które zależą również od rezystancji uziemienia stacji połączonych żyłami powrotnymi. Dane dla takich analiz można uzyskać prowadząc badania symulacyjne rozpływu prądów zwarć doziemnych dla aktualnego sposobu pracy punktu neutralnego sieci przy koniecznym uwzględnieniu rezystancji uziemień żył powrotnych, przez które przepływa prąd zwarciowy.

W przypadku mostów kablowych uziemienie obustronne żył powrotnych lub ich jednostronne odziemienie nie przyczynia się w sposób istotny do zmiany prądu przepływającego przez uziom stacji WN/SN, a zatem nie ma istotnego wpływu na zagrożenie porażeniowe podczas zwarć jednofazowych na tej stacji.

6.2. Przepięcia na izolowanych końcach żył powrotnych kabli

Odziemianie żył powrotnych nie powoduje negatywnych skutków przepięciowych podczas pracy normalnej. Zagrożenie stwarzają zwarcia doziemne oraz wielofazowe w kablach SN, gdzie jak wynika z przeprowadzonych symulacji obliczeniowych, wartość szczytowa przepięć ziemnozwarciowych w odziemionych żyłach powrotnych może być rzędu kilkunastu kV, co może stanowić zagrożenie przepięciowe dla izolacji kabla. Na ryzyko ewentualnego uszkodzenia powłoki wpływa nie tylko wartość szczytowa przepięć, ale również częstotliwość występowania zwarć w linii kablowej. Szczególnie częste narażenia mogą występować w ciągach kablowo napowietrznych, ponieważ można się spodziewać dużej ilości zwarć

w obszarze odcinków napowietrznych i negatywnego skumulowanego wpływu na izolację żył powrotnych odziemianego kabla [31, 39]. W przypadku zastosowania odizolowania dwóch żył powrotnych kabla, remedium na możliwość wzmacniania fali przepięciowej na krańcu kabla może być połączenie dwóch żył powrotnych na krańcach odizolowanych, co ogranicza zmianę impedancji falowej żyły powrotnej i powinno powodować ograniczenia wartości przepięć, lecz powoduje to wzrost strat w zwartych żyłach.

Obustronne uziemienie jednej lub dwóch żył powrotnych, jak wykazały badania modelowe, ogranicza napięcia w stanach ustalonych i nieustalonych na pozostałych odziemionych żyłach powrotnych w stosunku do napięć w stanach ustalonych i nieustalonych występujących w przypadku odziemienia trzech żył powrotnych. Można przypuszczać, że mechanizm oddziaływania obustronnie uziemionej żyły powrotnej jest zbliżony do oddziaływania ograniczającego przepięcia poprzez kabel ECC stosowany przy kablach wysokiego napięcia [51, 52].

Zastosowanie odziemiania żył powrotnych w przypadku linii kablowych wymaga przeprowadzenia dalszych badań pod względem możliwych przepięć podczas zwarć i analizy możliwych do zastosowania sposobów zapobiegania ich skutkom, jak również wpływu przepięć na izolację żyły powrotnej.

Wyniki symulacji wskazują, że przepięcia w mostach kablowych nawet podczas zwarć w ciągach zasilających cechują się stosunkowo niedużą wartością szczytową, ale należy podkreślić, że występują częściej tzn. podczas wszystkich zwarć w dowolnym miejscu w sieci zasilanej z danej stacji elektroenergetycznej. Teoretycznie mosty kablowe można bezpiecznie 1-stronnie całkowicie odziemić, ponieważ nawet w przypadku zwarć dwufazowych podwójnych wartość szczytowa przepięć, wynosi poniżej 2 kV i jest poniżej 5 kV wartości napięcia stałego próby napięciowej przeprowadzonej przy oddaniu kabla do eksploatacji.

7. Podsumowanie

Rosnący udział linii kablowych, wykonanych przy wykorzystaniu kabli jednożyłowych z żyłami powrotnymi, w całkowitej długości linii rozdzielczych średniego napięcia skłaniają do analiz optymalizacji sposobu eksploatacji tych kabli a w szczególności wyboru sposobu zapobiegania nadmiernym stratom w żyłach powrotnych tych kabli. W pracy dokonano analizy wartości strat przy obustronnym uziemieniu żył powrotnych dla linii i mostów kablowych w zależności od obciążenia i jego zmienności dla profilu obciążenia charakterystycznego dla odbiorców bytowo-komunalnych. Przeprowadzono symulacje komputerowe pracy odcinków kabla w funkcji obciążeń przenoszonych i osiągnięte wyniki zweryfikowano w zakresie strat w żyłach powrotnych przy wykorzystaniu pomiarów na obiektach rzeczywistych. Weryfikacja pomiarowa wskazała na 25% niedoszacowanie wartości prądów indukowanych w żyłach powrotnych otrzymywanych w wyniku symulacji obliczeniowych. Jedną z przyczyn większych strat może być zawartość wyższych harmonicznych prądów w żyłach kabla.

Przeprowadzone symulacje, z uwzględnieniem korekty wynikającej z pomiarów, mogą być podstawą do oceny strat finansowych ponoszonych przy obustronnym uziemieniu żył powrotnych kabli w okresie pracy kabla. Poziom strat przy obostronnym uziemieniu żył powrotnych może skłaniać do modyfikacji sposobu pracy tych żył poprzez zastosowanie środków zaradczych takich jak przeplecenia, odziemienie jednostronne żył powrotnych czy zastosowanie mniejszego przekroju tych żył. W artykule zaproponowano uproszczoną metodykę oceny ekonomicznej redukcji tych strat podając zależności na obliczenie oszczędności na stratach w postaci wartości zdyskontowanej na rok analizy tych oszczędności

za okres 25 lat przy założeniu określonego wzrostu cen energii oraz wzrostu obciążenia w analizowanym okresie. Wyznaczone oszczędności stanowią podstawę do podjęcia decyzji o zastosowaniu modyfikacji sposobu pracy czy też przekroju żył powrotnych.

Podstawowym wnioskiem z przeprowadzonych symulacji jest zalecenie ograniczenia przekroju żył powrotnych w kablach do przekroju 25 mm^2. Przy obecnym poziomie obciążenia linii kablowych w sieci rozdzielczej zasilającej odbiorców bytowo komunalnych poziom strat nie uzasadnia zastosowania dość kosztownego środka zaradczego w postaci przepleceń żył powrotnych kabli.

Stosunkowo prostym środkiem zaradczym jest odziemianie żył powrotnych na jednym z końców. Zastosowanie tego rozwiązania poprzez jednostronne odziemienie wszystkich żył jest możliwe w przypadku mostów kablowych zbudowanych z kabli jednożyłowych. Dla linii kablowych, w przypadkach wykazujących opłacalność ograniczania strat poprzez odziemianie, można rozważyć odziemienie jednej lub dwóch żył powrotnych pozostawiając uziemioną obustronnie jedną żyłę dla zapewnienia możliwości redukcji prądu zwarciowego poprzez uziom stacji.

W artykule przedstawiono analizę zagadnienia ograniczenia prądu zwarcia jednofazowego w następstwie odziemiania żył powrotnych. Przedstawiono wyniki redukcji prądu spływającego do ziemi poprzez uziom na podstawie zależności analitycznych, symulacji komputerowych oraz badań kabli ułożonych w ziemi przy ich zasileniu napięciem obniżonym. Procedury wyznaczania rozpływu prądów zwarcia powinny uwzględniać rezystancje własne żył powrotnych oraz rezystancje uziemień stacji, do których przyłączane są żyły powrotne. Na terenach miejskich gdzie można uznać, że systemy uziemiające tworzą zintegrowane instalacje uziemiające, zagrożenia porażeniowe nie powinny być przeszkodą w realizacji odziemienia dla ograniczenia strat.

Przeprowadzono symulacje komputerowe zagrożeń przepięciami izolacji powłoki kabla na odziemionych żyłach powrotnych mogące wystąpić podczas zwarć. Zaproponowano dla celów dalszych badań sposoby mogące ograniczać wartości takich przepięć. Przy dwóch żyłach odziemionych proponuje się zwarcie tych żył na odziemionym końcu z powodu braku zmiany impedancji tych żył dla fal przepięciowych, lecz za cenę podniesienia poziomu strat jak dla przypadku zastosowania tylko jednej żyły odziemionej. Alternatywą jest pozostawienie dwóch uziemionych żył powrotnych dla wzmocnienia efektu uzyskiwanego poprzez przewody ECC w kablach wysokiego napięcia.

Literatura

1. Arritt R, Dugan R. Comparing Load Estimation Methods for Distribution System Analysis. 22nd International Conference and Exhibition on Electricity Distribution. CIRED 2013.
2. Barrett J S, Anders G J. Circulating current and hysteresis losses in screens, sheaths and armour of electric power cables – mathematical models and comparison with IEC Standard 287. IEE Proceedings - Science, Measurement and Technology 1997; 144: 101-110.
3. Becker J, Musique F. Economical 36 kV Cable System for the Belgian Network. 16th International Conference and Exhibition on Electricity Distribution. CIRED 2001.
4. CEER Benchmarking Report 6.1 on the Continuity of Electricity and Gas Supply. Council of European Energy Regulators asbl 2018.

5. Charlton T E, Hocaoglu M H, Karacasu O, Marican A M A. Impact of cable sheath sizing, material and connections upon the safety of electrical power installations. 21st International Conference on Electricity Distribution. CIRED 2011.
6. Czapp S, Dobrzyński K, Klucznik J, Lubośny Z. Analiza napięć indukowanych w żyłach powrotnych kabli wysokiego napięcia dla ich wybranych konfiguracji. XVIII Międzynarodowa Konferencja Aktualne Problemy w Elektroenergetyce APE 2017.
7. Deschamps P, Toravel Y. Reduction of Technical and Non-Technical Losses in Distribution Networks. CIRED Overview, final report 2017.
8. DIgSILENT Power Factory 2018 Technical Reference Documentation. Cable System, ElmCabsys, TypCabsys.
9. DIgSILENT Power Factory 2018 User Manual.
10. Dong X, Yang Y, Zhou C, Hepburn D M. Online Monitoring and Diagnosis of HV Cable Faults by Sheath System Currents. IEEE Transactions on Power Delivery 2017; 32: 2281-2290.
11. Duda D, Szadkowski M. Ochrona przeciwprzepięciowa osłon kabli WN w różnych układach połączeń żył powrotnych. Przegląd Elektrotechniczny 2014; 10: 37-40.
12. Elektroenergetyczne linie kablowe średniego napięcia. Standard w sieci dystrybucyjnej Enea Operator Sp. z o.o. 2017.
13. European Standard PN-EN 50341-1:2013-03 Overhead electrical lines exceeding AC 1 kV - Part 1: General requirements - Common specifications.
14. European Standard PN-EN 50522:2011 Earthing of power installations exceeding 1 kV a.c.
15. European Standard PN-EN 61936-1:2011 Power installations exceeding 1 kV a.c. - Part 1: Common rules.
16. European Standard PN-HD 60364-4-442:2012 Electrical Installations of Buildings - Protection for Safety - Protection of Low-Voltage Installations Against Temporary Overvoltages and Faults Between High-Voltage Systems and Earth.
17. Gołaś A, Ciesielka W, Szopa K, Zydroń P, Bąchorek W, Benesz M, Kot A, Moskwa S. Analysis of the possibilities to improve the reliability of a 15 kV overhead line exposed to catastrophic icing in Poland. Eksploatacja i Niezawodność – Maintenance and Reliability 2019; 21(2): 282–288.
18. Gouda O E, Farag A A. Factors affecting the sheath losses in single-core underground power cables with two-points bonding method. International Journal of Electrical and Computer Engineering (IJECE) 2012; 2: 7–16.
19. Gouramanis K V, Kaloudas Ch G, Papadopoulos T A, Papagiannis G K, Stasinos K. Sheath Voltage Calculations in Long Medium Voltage Power Cables. IEEE Trondheim PowerTech 2011.
20. Heiss A, Balzer G, Schmitt O, Richter B. Surge arresters for cable sheath preventing power losses in MV networks. 16th International Conference and Exhibition on Electricity Distribution. CIRED 2001.
21. IEEE Guide for the Application of Sheath-Bonding Methods for Single-Conductor Cables and the Calculation of Induced Voltages and Currents in Cable Sheaths. IEEE Standard 575 1988.
22. Instrukcja Ruchu i Eksploatacji Sieci Dystrybucyjnej Enea Operator Sp. z o.o.: Standardowe profile zużycia energii na rok 2017. http://www.operator.enea.pl/21/instrukcje/instrukcje-iriesd-883.html, dostęp: 24.07.2019.

23. Jensen C F, Nanayakkara O M K K, Rajapakse A D, Gudmundsdottir U S, Bak C L. Online fault location on AC cables in underground transmissions systems using sheath currents. Electric Power Systems Research 2014; 115: 74-79.
24. Jung C K, Lee J B, Kang J W, Wang X H, Song Y H. Sheath Current Characteristic and Its Reduction on Underground Power Cable Systems. IEEE Power Engineering Society General Meeting 2005.
25. Kauppinen M, Pylvanainen J, Karjalainen J, Sihvola V, Experiences of using AMI system for DSO's business operation. CIRED – Open Access Proceedings Journal 2017; 1: 2756-2759.
26. Korab R, Siwy E. Statistical analysis of the double line-to-ground short-circuit current in MV urban network for the power cable metallic screen rating. International Conference on Probabilistic Methods Applied to Power Systems 2006.
27. Korab R, Siwy E, Żmuda K. Analiza możliwości redukcji przekroju żył powrotnych w kablach średniego napięcia w sieciach miejskich. Zeszyty Naukowe Politechniki Śląskiej. Elektryka 2004; 189: 111-120.
28. Korab R, Siwy E, Żmuda K. Podstawy racjonalnego doboru żył powrotnych (ekranów) kabli 6-20 kV różnego typu w sieciach miejskich. Elektroenergetyczne linie kablowe 2004; 148-149: 19-28.
29. Kornatka M. Analysis of the exploitation failure rate in Polish MV networks. Eksploatacja i Niezawodność – Maintenance and Reliability 2018; 20(3): 413–419.
30. Krawiec H. Przyczyny grzania się bednarki i żył powrotnych kabli 6 kV. Zeszyty Problemowe – Maszyny Elektryczne 2014; 1: 141-145.
31. Li J, Xu L, Chen X, Zhao A, Liu J, Zhao X, Deng J, Zhang G. Analysis of Statistical and Frequency Characteristics of Transient Overvoltage of Hybrid Cable-OHL Lines. China International Conference on Electricity Distribution (CICED) 2018; 2650-2654.
32. Lin Y, Xu Z. Cable Sheath Loss Reduction Strategy Research Based on the Coupled Line Model. IEEE Transactions on Power Delivery 2015; 30: 2303-2311.
33. Lin Y, Yang F, Xu Z, Weng H. Cable Sheath Loss Analysis Based on Coupled Line Model. International Conference on Power System Technology 2014.
34. Łowczowski K. Badanie wpływu ułożenia kabli na straty energii w żyle powrotnej – symulacja w programie PowerFactory. Przegląd Elektrotechniczny 2016; 10: 54-57.
35. Moore G F. Electric Cables Handbook, Third Edition. Oxford: Blackwell Science Ltd, 1997.
36. Niewiedział R, Niewiedział E, Wyznaczanie czasu trwania strat maksymalnych w sieciach elektroenergetycznych modelami obliczeniowymi. http://www.sep.krakow.pl/nbiuletyn/nr57ar2.pdf, dostęp 5.07.2019.
37. Novak B, Koller L, Berta I. Loss reduction in cable sheathing. Renewable Energy and Power Quality Journal 2010; 1: 293-297.
38. Novak B, Koller L. Current distribution and losses of grouped underground cables. IEEE Transactions on Power Delivery 2011; 26: 1515-1521.
39. Orsagova J, Toman P. Transient overvoltages on distribution underground cable inserted in overhead line. 20th International Conference on Electricity Distribution. CIRED 2009.
40. Polityka Energetyczna Polski do 2040 roku, projekt. Warszawa: Ministerstwo Energii, 2018.

41. Pons E, Colella P, Napoli R, Tommasini R. Impact of MV Ground Fault Current Distribution on Global Earthing Systems. IEEE Transactions on Industry Applications 2015; 51: 4961-4968.
42. Popović L M. Ground Fault Current Distribution When a Ground Fault Occurs in HV Substations Located in an Urban Area. Progress In Electromagnetics Research B 2014; 59: 167-179.
43. Rakowska A. Service Experiences for MV Cable Network - Optimistic or Pessimistic State of The Art. Jicable Conference: The leading forum about Insulated Power Cables 2007.
44. Riba Ruiz J R, Garcia A, Alabern Morera X. Circulating sheath currents in flat formation underground power lines. 2007.
45. Rozporządzenie Rady Ministrów z dnia 10.12.2010 r. w sprawie Klasyfikacji Środków Trwałych (KŚT).
46. Sakalkale S, Kanakgiri K. Study of Underground Power Cable Considering Sheath Circulating Current. Proceedings of Fourth IRF International Conference 2014; 7-10.
47. Sapienza G, Noce C, Valvo G. Network Technical Losses Precise Evaluation Using Distribution Management System and Accurate Network Data. 23rd International Conference on Electricity Distribution. CIRED 2015.
48. Sobral A, Moura A, Carvalho M. Technical Implementation of Cross Bonding on Underground High Voltage Lines Projects. 21st International Conference on Electricity Distribution. CIRED 2011.
49. Sullivan R G, Abplanalp J M, Lahti S, Beckman K J, Cantwell B L, Richmond P. Electric Transmission Visibility and Visual Contrast Threshold Distances in Western Landscapes. National Association of Environmental Professionals Annual Conference 2014; 1-46.
50. Tarko R, Benesz M, Nowak W, Szpyra W. Statystyczna analiza zakłóceń zwarciowych dla określenia przekroju żył powrotnych kabli średnich napięć. Przegląd Elektrotechniczny 2016; 7: 186-189.
51. Todorovski M, Ackovski R. Equivalent Circuit of Single-Core Cable Lines Suitable for Grounding System Analysis Under Line-to-Ground Faults. IEEE Transactions on Power Delivery 2014; 29: 751-759.
52. Żmuda K. Elektroenergetyczne układy przesyłowe i rozdzielcze. Wybrane zagadnienia z przykładami. Gliwice: Wydawnictwo Politechniki Śląskiej, 2016.

dr hab. inż. Stanisław Młynarski
Politechnika Krakowska
31-864 Kraków, al. Jana Pawła II 37,
tel.: (012) 374 33 22,
email: mlynarski_st@poczta.onet.pl;

dr hab. inż. Robert Pilch
AGH Akademia Górniczo-Hutnicza w Krakowie
30-059 Kraków, al. Mickiewicza 30,
tel.: (012) 617 31 10,
email: pilch@agh.edu.pl;

dr inż. Maksymilian Smolnik
AGH Akademia Górniczo-Hutnicza w Krakowie
30-059 Kraków, al. Mickiewicza 30,
tel.: (012) 617 31 25,
email: smolnik@agh.edu.pl;

prof. Jan Szybka
AGH Akademia Górniczo-Hutnicza w Krakowie
30-059 Kraków, al. Mickiewicza 30,
tel.: (012) 617 31 10,
email: szybja@agh.edu.pl

mgr inż. Grzegorz Wiązania
AGH Akademia Górniczo-Hutnicza w Krakowie
30-059 Kraków, al. Mickiewicza 30,
tel.: (012) 617 31 25,
email: wiazania@agh.edu.pl

Model adaptacyjnej strategii prewencyjnej odnowy złożonych obiektów technicznych

Słowa kluczowe: *odnowa profilaktyczna; symulacyjna ocena efektywności; inżynieria niezawodności*

Streszczenie: W artykule przedstawiono wyniki analizy opracowanych modeli do planowania odnowy profilaktycznej złożonych obiektów technicznych, które oparto o dwa różniące się od siebie zestawy założeń. Rozwiązywany problem dotyczy określenia wspólnego czasu odnowy profilaktycznej grupy części lub podzespołów złożonego obiektu. Pierwszy z opracowanych modeli (model planowej strategii odnowy prewencyjnej) pozwala określić zasadność przeprowadzenia ustalonego wcześniej, planowego odnowienia prewencyjnego części obiektu, która została już odnowiona poawaryjnie. Drugi model (model adaptacyjnej strategii odnowy prewencyjnej) umożliwia wyznaczenie najbliższego wspólnego czasu odnowy profilaktycznej grupy części, z których jedna aktualnie podlega odnowie poawaryjnej. Początkowe (wyjściowe) strategie odnowy profilaktycznej każdej części bądź podzespołu wyznaczone zostały za pomocą standardowych narzędzi do planowania odnawiania profilaktycznego (modeli decyzyjno-losowych wykorzystujących programowanie dynamiczne Bellmana). Posługując się opracowanymi modelami odnowy, przeprowadzono przykładowe obliczenia symulacyjne, których

wyniki przedstawiono w postaci całkowitych kosztów obsługiwania dla każdej z uzyskanych strategii. Przedmiotem analizy były wybrane cechy geometryczne koła pojazdu szynowego, których wartości zmieniają się na skutek zużycia w procesie eksploatacji. Na podstawie tego rodzaju analiz można wybrać lepszy (tj. efektywniejszy ekonomicznie) z modeli dla konkretnego zastosowania w praktyce.

Wykaz oznaczeń:

t_i^* – czas odnowy prewencyjnej *i*-tego elementu układu,
k_{ai} – koszt awarii *i*-tego elementu układu,
k_{oi} – koszt odnowy prewencyjnej *i*-tego elementu układu,
t_{max}^* – największa wartość czasu odnowy prewencyjnej jednego spośród wszystkich elementów układu,
t_{min}^* – najmniejsza wartość czasu odnowy prewencyjnej jednego spośród wszystkich elementów układu,
t_{ai} – czas wystąpienia awarii *i*-tego elementu,
t^* – wspólny czas pracy do odnowy prewencyjnej elementów wyznaczony przed awarią *i*-tego elementu,
$F_i(t)$ – wartość dystrybuanty rozkładu czasu pracy do uszkodzenia *i*-tego elementu,
k_{ai} – koszt awarii *i*-tego elementu układu,
k_{oi} – koszt odnowy prewencyjnej *i*-tego elementu układu,
$D(t_{ai})$ – wskaźnik decyzyjny zasadności odnowy,
T_h – skończony horyzont czasowy eksploatacji obiektu,
T_{hmax} – maksymalny czas eksploatacji bez odnowy prewencyjnej lub przeglądu,
Δt – długość przedziału czasu w obliczeniach w modelu decyzyjno-losowym,
n_{OT} – liczba działań obsługowych w rozważanym horyzoncie czasowym,
n – liczba powtórzeń symulacji,
dt – krok czasowy w symulacji,
t_U – wiek układu od początku symulacji,
t_{Ei} – wiek i-tego elementu od początku symulacji,
Pu_{Ei} – prawdopodobieństwo uszkodzenia i-tego elementu w symulacji,
RND_{Ei} – liczba losowa z przedziału [0,1], z rozkładu jednostajnego, dla i-tego elementu,
t_{O_opt} – optymalny okres odnowy układu w symulacji.

1. Wprowadzenie

Zapewnienie efektywnej ekonomicznie i bezpiecznej eksploatacji złożonych obiektów technicznych wymaga stosowania odpowiedniej strategii ich obsługiwania [4]. Jakość procesu eksploatacji, opisywana różnymi wskaźnikami [19], w znacznej mierze zależy od prawidłowej odnowy obiektów technicznych. Zagadnienie to jest szczególnie ważne, gdy dotyczy układów, w których uszkodzenia elementów mogą doprowadzić do wystąpienia zagrożenia zdrowia i życia ludzi lub bardzo dużych strat ekonomicznych. W tego typu przypadkach może być uzasadnione odnawianie prewencyjne elementów układu, a optymalny czas jego przeprowadzania wyznacza się przy wykorzystaniu różnych modeli odnów prewencyjnych [15, 20]. W tym celu wykorzystuje się dane ekonomiczne związane z awariami i odnowami oraz charakterystyki niezawodnościowe elementów układu. Oprócz klasycznych modeli odnawiania, w literaturze dostępnych jest wiele modeli odnowy [10] ukierunkowanych na zapewnienie wymaganego poziomu niezawodności całego układu [12, 21], niektóre z nich opierają się na zastosowaniu metod symulacji [5, 8]. Stosowane są również modele

uwzględniające możliwość wykonania odnowy częściowej (por. [7, 8]), czy też wykorzystujące dane z dodatkowych kontroli stanu technicznego obiektu, jeżeli jest możliwość przeprowadzania ich w trakcie pracy układu [2, 3]. Szczegółową i wyczerpującą klasyfikację istniejących modeli odnawiania prewencyjnego zamieszczono w opracowaniu [20].

Model wykorzystany do wyznaczenia optymalnego czasu odnowy prewencyjnej (t^*) powinien być dostosowany do specyfiki i warunków eksploatacji rozważanego obiektu. Tylko wtedy, stosowanie wyznaczonej z jego użyciem strategii odnów prewencyjnych, będzie uzasadnione ekonomicznie i zapewni zmniejszenie awaryjności obiektu [11, 14].

Celem artykułu jest opracowanie adaptacyjnego modelu odnawiania złożonych obiektów technicznych oraz symulacyjnej metody do oceny efektywności stosowania wyznaczonej tym modelem strategii odnowy.

W artykule rozważano specyficzny problem odnowy obiektu złożonego, którego części funkcjonują w strukturze szeregowej i ich uszkodzenia traktuje się jako niezależne od siebie, jednak ze względów organizacyjnych i ekonomicznych indywidualna odnowa każdej z nich może być mniej opłacalna niż odnawianie blokowe.

W skomplikowanych przypadkach, gdy odnowy pewnych części analizowanych złożonych obiektów technicznych związane są czasowo z działaniami obsługowymi dotyczącymi innych jednostek lub układów tego samego obiektu, po wystąpieniu awarii konieczne jest rozważenie modyfikacji zaplanowanej strategii odnowy uszkodzonej części lub nawet całego obiektu. Problem ten jest przedmiotem rozważań podjętych w niniejszym opracowaniu.

Opracowane rozwiązania stanowią istotne uzupełnienie istniejących modeli odnowy ze względu na potrzeby występujące w praktyce.

W artykule przedstawiono opracowaną oryginalną metodę oceny efektywności stosowania proponowanych strategii przeprowadzoną przy użyciu autorskiego programu komputerowego. Opracowaną metodę zweryfikowano przy wykorzystaniu przykładu praktycznego dotyczącego uszkodzeń kół pojazdu szynowego typu elektryczny zespół trakcyjny.

2. Modele odnawiania prewencyjnego złożonych układów

Obiekty techniczne składają się z wielu elementów konstrukcyjnych, które posiadają indywidualne charakterystyki eksploatacyjne. Empiryczne wskaźniki niezawodności części składowych opisywane są często różnymi modelami i wykazują znacząco różne przebiegi dla każdej z nich [7]. Powoduje to problemy decyzyjne dotyczące wyznaczania terminów i zakresu dokonywanej ze względów bezpieczeństwa odnowy całych obiektów lub ich zespołów i części, a w konsekwencji prowadzenia ich racjonalnie i ekonomicznie uzasadnionej eksploatacji [por. 6, 16].

2.1. Wyznaczanie wspólnego czasu odnowy profilaktycznej elementów układu

W przypadku odnawiania prewencyjnego układów, które składają się z pewnej liczby różniących się od siebie elementów, ze względów organizacyjnych i ekonomicznych najczęściej poszukuje się wspólnego czasu przeprowadzenia odnowy dla wszystkich elementów – tj. dla całego układu [17]. Jeśli w wykorzystywanym modelu odnawiania prewencyjnego pewien zespół złożonego obiektu technicznego rozważany jest jako całość – istnieje określona dla niego funkcja niezawodności i koszty odnowy oraz awarii, to wyznaczony

optymalny czas odnowy odnosi się do całego tego zespołu i jest on odnawiany prewencyjnie i poawaryjnie zawsze jako całość (odnowie podlegają jednocześnie wszystkie jego części).

Problem wyznaczenia wspólnego czasu odnowy pojawia się, gdy dla poszczególnych elementów układu wyznaczone są optymalne (według przyjętego modelu odnawiania prewencyjnego), różniące się od siebie (ale zbliżone) czasy odnawiania prewencyjnego. Proponowany sposób ujednolicenia czasu odnowy prewencyjnej zakłada uwzględnienie indywidualnych czasów odnowy elementów (t_i^*), ich kosztów odnowy prewencyjnej (k_{oi}), kosztów awarii (k_{ai}) oraz przyrostów wartości dystrybuant w zakresie rozstępu indywidualnych czasów odnów prewencyjnych dla wszystkich rozważanych elementów. Przyrost wartości dystrybuanty oraz koszty odnowy i wystąpienia awarii będą dla każdego z elementów stanowić wagę jego indywidualnego czasu odnowy prewencyjnej (t_i^*). Wspólny czas odnowy (t^*) proponuje się wyznaczać wg zależności:

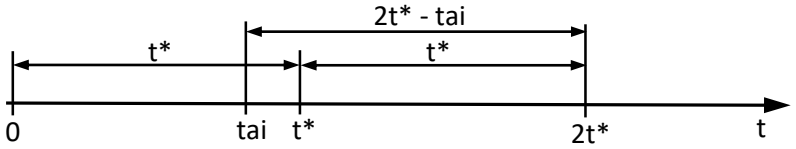
$$t^* = \frac{\sum_{i=1}^{n} t_i^* \cdot (k_{ai}+k_{oi}) \cdot [F_i(t_{max}^*) - F_i(t_{min}^*)]}{\sum_{i=1}^{n} (k_{ai}+k_{oi}) \cdot [F_i(t_{max}^*) - F_i(t_{min}^*)]}, \tag{1}$$

gdzie:
t_i^* – czas odnowy prewencyjnej i-tego elementu układu,
k_{ai} – koszt awarii i-tego elementu układu,
k_{oi} – koszt odnowy prewencyjnej i-tego elementu układu,
t_{max}^* – największa wartość czasu odnowy prewencyjnej jednego spośród wszystkich elementów układu,
t_{min}^* – najmniejsza wartość czasu odnowy prewencyjnej jednego spośród wszystkich elementów układu.

Wspólny czas odnowy prewencyjnej rozpatrywanego zespołu (t^*) może być również wyznaczany z uwzględnieniem ustalonych cyklicznych przeglądów i działań obsługowych realizowanych dla innych zespołów i części tego samego złożonego obiektu [9]. Takie podejście jest często stosowane w praktyce eksploatacyjnej w dużych przedsiębiorstwach transportowych, gdzie wynikające z zastosowanego modelu odnawiania prewencyjnego optymalne okresy odnowy i przeglądów poszczególnych zespołów są skracane lub wydłużane tak, aby stanowiły całkowitą wielokrotność przeprowadzanych zgodnie z harmonogramem czasowym czynności obsługowych większej liczby zespołów i części tego obiektu. Działanie takie pozwala zmniejszyć liczbę przestojów złożonego obiektu technicznego przynosząc wymierne efekty ekonomiczne.

2.2. Model planowej strategii odnowy prewencyjnej układu

Po wyznaczeniu wspólnego czasu odnowy prewencyjnej dla wszystkich elementów układu (t^*), strategia wspólnej cyklicznej ich odnowy po przepracowaniu tego czasu może być łatwo stosowana do chwili, gdy któryś z elementów ulegnie awarii. Gdy takie zdarzenie wystąpi pojawia się problem decyzyjny, czy odnowiony po awarii element odnowić również w zaplanowanym wspólnym terminie (mimo, że nie osiągnie on jeszcze wtedy czasu pracy t^* do zaplanowanej odnowy), czy pozostawić go do kolejnego wspólnego terminu odnowy (Rys. 1), czyli wydłużyć czas jego pracy do odnowienia prewencyjnego ponosząc zwiększone ryzyko wystąpienia awarii.



Rys. 1. Planowy czas pracy elementu do odnowy prewencyjnej i wydłużony po wystąpieniu uszkodzenia

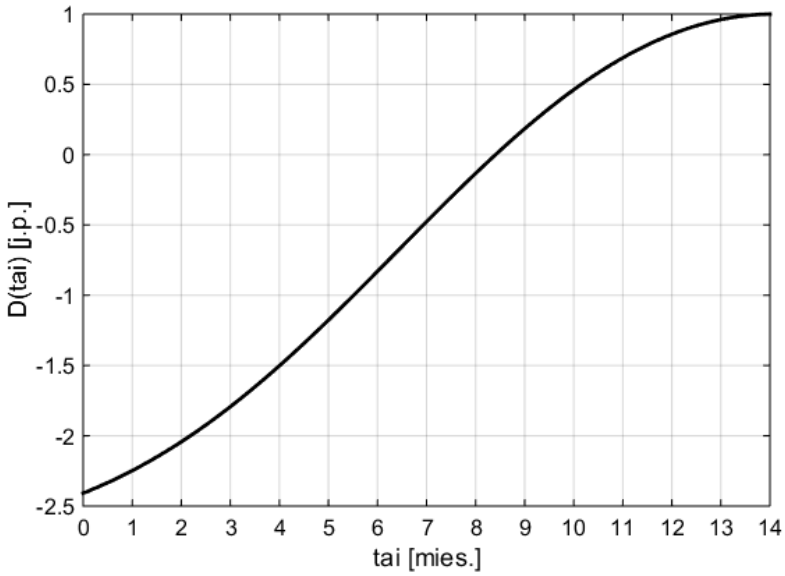
Rozwiązanie problemu wymaga uwzględnienia zwiększonego prawdopodobieństwa uszkodzenia się tego elementu w wydłużonym okresie eksploatacji ($2t^* - t_{ai}$) jak i kosztów związanych z jego ewentualnym dodatkowym odnowieniem (k_{oi}), gdy odnowa elementu zostałaby przeprowadzona w ramach odnowy całego układu mimo, że element ten nie osiągnie jeszcze planowego czasu pracy t^* do swojej odnowy prewencyjnej (przepracuje tylko czas $t^* - t_{ai}$). Wskaźnik na podstawie którego można w takiej sytuacji podejmować decyzję, w stosunku do uszkodzonego elementu, proponuje się wyznaczać z zależności:

$$D(t_{ai}) = k_{oi} + F_i(t^* - t_{ai}) \cdot k_{ai} - \frac{F_i(2 \cdot t^* - t_{ai}) - F_i(t^*)}{1 - F_i(t^*)} \cdot k_{ai}, \tag{2}$$

gdzie:
t_{ai} – czas wystąpienia awarii i-tego elementu,
t^* – wspólny czas pracy do odnowy prewencyjnej elementów wyznaczony przed awarią i-tego elementu,
$F_i(t)$ – wartość dystrybuanty rozkładu czasu pracy do uszkodzenia i-tego elementu,
k_{ai} – koszt awarii i-tego elementu układu,
k_{oi} – koszt odnowy prewencyjnej i-tego elementu układu.

Jeśli wartość $D(t_{ai}) \leq 0$, to uzasadniona jest odnowa elementu wraz z pozostałymi mimo, że od odnowy poawaryjnej w chwili t_{ai} nie przepracował on jeszcze czasu równego t^*. Gdy $D(t_{ai}) > 0$, to uzasadnione jest wydłużenie czasu pracy elementu i nie odnawianie go przy najbliższej odnowie pozostałych elementów, lecz dokonanie tego dopiero w trakcie kolejnej odnowy układu, po najbliższej wspólnej odnowie jego elementów.

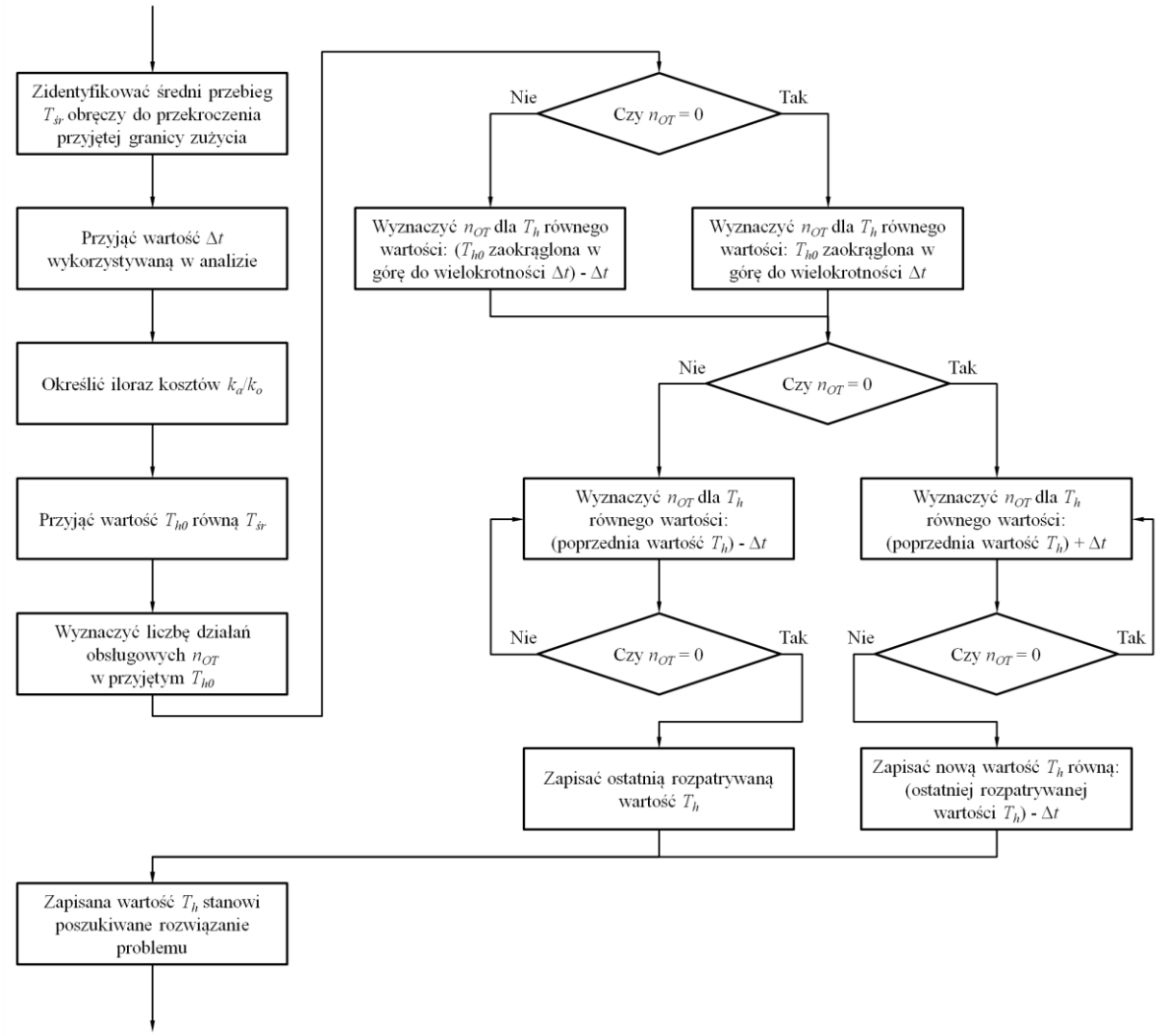
Przy znanych wartościach kosztów awarii i odnowy oraz dystrybuanty czasów pracy do uszkodzenia poszczególnych elementów, dla każdego z nich można wyznaczyć graniczną wartość czasu $t_{ai} = t_{i\_gr}$. Jest to wartość czasu, dla którego $D(t_{ai}) = 0$. Jeśli awaria wystąpi przed osiągnięciem przez element tej wartości czasu pracy, to uzasadniona jest jego dodatkowa odnowa wraz z pozostałymi elementami w najbliższym wspólnym terminie odnowy prewencyjnej. Jeśli awaria elementu wystąpi już po osiągnięciu tego czasu pracy, uzasadnione jest wydłużenie czasu pracy elementu bez odnowy. Przykład przebiegu funkcji $D(t_{ai})$ przedstawiono na rysunku 2. Krzywą wykreślono zgodnie z równaniem (2) dla elementu o normalnym rozkładzie czasu pracy do uszkodzenia N(m = 21,5 [mies.]; σ = 4,75 [mies.]), przy ilorazie kosztu awarii i kosztu odnowy ka/ko = 4 i wyznaczonym okresie odnowy prewencyjnej t^* =14 [mies.].



Rys. 2. Przykład przebiegu funkcji $D(t_{ai})$

Z wykresu można odczytać wyznaczoną w tym przypadku wartość t_{i_gr} wynoszącą 8,5 [mies.]. Oznacza to, że jeśli uszkodzenie elementu wystąpi do 8,5 miesiąca eksploatacji, uzasadniona jest jego dodatkowa odnowa prewencyjna z innymi elementami zaplanowana w 14 miesiącu. Jeśli awaria wystąpi już po 8,5 miesiąca, okres jego eksploatacji można wydłużyć (o maksymalnie 5,5 miesiąca) i odnowić ten element nie przy najbliższej wspólnej odnowie elementów układu, ale dopiero w ramach kolejnej.

Innym przedstawionym w literaturze [13] sposobem wyznaczenia wartości czasu t_{i_gr} jest wykorzystanie modeli decyzyjno-losowych opartych na programowaniu dynamicznym i zasadzie optymalności Bellmana. Modele te umożliwiają obliczenie optymalnego okresu odnowy prewencyjnej (t^*) w skończonym horyzoncie czasowym (T_h) eksploatacji obiektu technicznego, ale przy ich zastosowaniu możliwe jest również obliczenie największej wartości T_h, dla której model ten wskazuje na brak opłacalności przeprowadzania odnowy prewencyjnej lub kontroli stanu technicznego układu [13], co pokazano z pomocą schematu na rysunku 3 [18]. Wartość tego maksymalnego czasu (T_{hmax}, bez odnowy prewencyjnej lub przeglądu) jest odpowiednikiem czasu $2t^* - t_{i\_gr}$ w modelu proponowanym w niniejszym opracowaniu. Znając T_{hmax} z modelu decyzyjno-losowego można również obliczyć $t_{i\_gr} = 2t^* - T_{hmax}$.



Rys. 3. Metoda wyznaczania największego przebiegu obręczy zestawu kołowego, dla którego rozpatrywane działania obsługowe pozostają nieopłacalne [18]

Cechą charakterystyczną przedstawionej planowej strategii odnowy jest fakt, że nie zmieniają się ustalone początkowo wspólne chwile odnowy prewencyjnej wszystkich elementów, wykonywanej w stałych odstępach czasu o wartości t^*. Stosowanie takiej strategii może mieć szczególne uzasadnienie, gdy odnowy rozważanego układu przeprowadzane są w stałych odstępach czasu t^* i w jego wielokrotnościach, związanych z innymi zaplanowanymi działaniami obsługowymi, całego obiektu lub pewnej ich zbiorowości.

2.3. Model adaptacyjnej strategii odnowy prewencyjnej układu

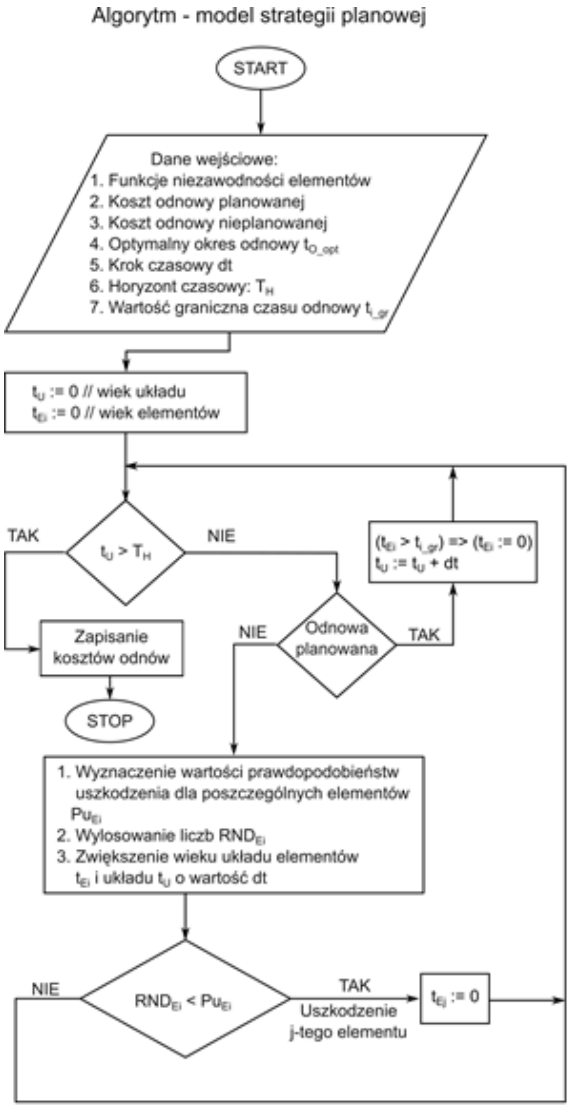
Drugą proponowaną możliwością odnawiania jest stosowanie strategii adaptacyjnej, w której możliwe jest modyfikowanie odstępu między chwilami wykonywania kolejnych odnów prewencyjnych całego układu. Odstęp ten wyznaczany jest na nowo po każdym wystąpieniu awarii dowolnego elementu układu. W modelu tym sytuacja wyjściowa jest taka sama jak w modelu planowej strategii, czyli istnieje wyznaczony stały i wspólny dla wszystkich elementów układu czas pracy do odnowy prewencyjnej (t^*). Strategia zgodna z tym czasem jest stosowana do chwili, gdy któryś z elementów uszkodzi się przed osiągnięciem wieku t^*. Gdy zdarzenie takie wystąpi, w strategii adaptacyjnej, następuje odnowa poawaryjna uszkodzonego elementu

i przy wykorzystaniu wiedzy, że pozostałe elementy nie uległy do tej chwili uszkodzeniu (modyfikacja rozkładów prawdopodobieństwa czasów pracy do uszkodzenia), następuje wyznaczenie kolejnego, wspólnego dla wszystkich elementów, czasu wykonania odnowy prewencyjnej.

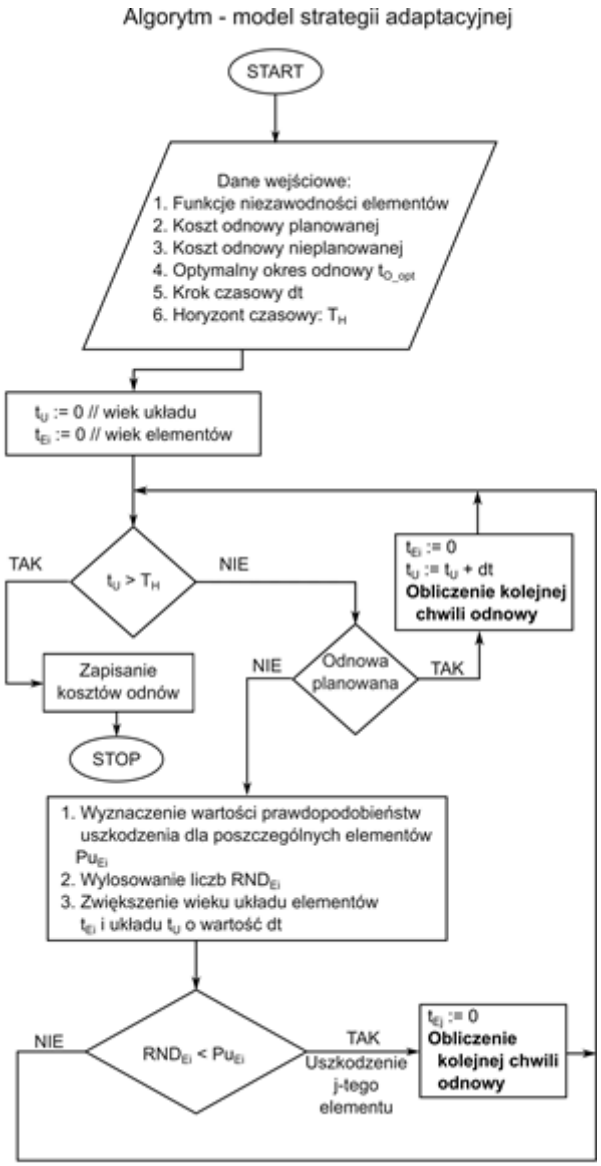
Zaproponowane dwie strategie odnawiania różnią się od siebie, a o tym, która w danym przypadku praktycznym będzie bardziej efektywna, można stwierdzić poprzez wyznaczenie wartości wskaźników ekonomicznych, wynikających z ich stosowania w założonym horyzoncie czasowym [1].

3. Model symulacyjny planowej i adaptacyjnej strategii odnawiania

Symulacyjny model obliczeniowy opracowano na podstawie założeń sformułowanych dla modelu strategii planowej i modelu strategii adaptacyjnej scharakteryzowanych w rozdziale 2. Opracowane algorytmy obliczeniowe pozwalają na wykonanie symulacji komputerowych pracy układu przy stosowaniu każdej z nich. Uproszczone algorytmy przedstawiające pojedynczą iterację symulacji pokazano na rysunkach 4a i 4b.



Rys. 4a. Uproszczony algorytm pojedynczej iteracji symulacji pracy układu według modelu strategii planowej



Rys. 4b. Uproszczony algorytmy pojedynczej iteracji symulacji pracy układu według modelu strategii adaptacyjnej

Szczegółową charakterystykę metody symulacyjnej wykorzystanej do obliczeń można znaleźć w [7], gdzie pokazano możliwość oszacowania niezawodności złożonego obiektu technicznego podlegającego dekompozycji. Na potrzeby niniejszego artykułu metoda ta stanowi bazę, która została w istotnym stopniu uzupełniona o nowe funkcje związane z odnowami prewencyjnymi badanych obiektów. Dzięki temu, poza szacowaniem niezawodności obiektu, możliwa jest również ocena efektywności ekonomicznej różnych strategii odnawiania prewencyjnego jego części.

Dla algorytmu bazującego na modelu strategii planowej przyjęto założenia:

- w przypadku uszkodzenia elementu następuje jego odnowa bez odnowy elementów nieuszkodzonych,
- element ten zostanie odnowiony w ramach odnowy profilaktycznej dla czasu t^*, jeżeli przepracował czas mniejszy niż t_{i_gr}.

Danymi wejściowymi do programu opracowanego na bazie tego algorytmu są: funkcje niezawodności elementów, koszt odnowy prewencyjnej dla każdego elementu k_{oi}, koszt odnowy poawaryjnej k_{ai}, optymalny okres odnowy elementów układu t^*, wartość graniczna czasu t_{i_gr}, krok czasowy dt oraz horyzont czasowy T_H symulacji i liczba symulacji n.

Dla algorytmu bazującego na modelu strategii adaptacyjnej przyjęto założenia:

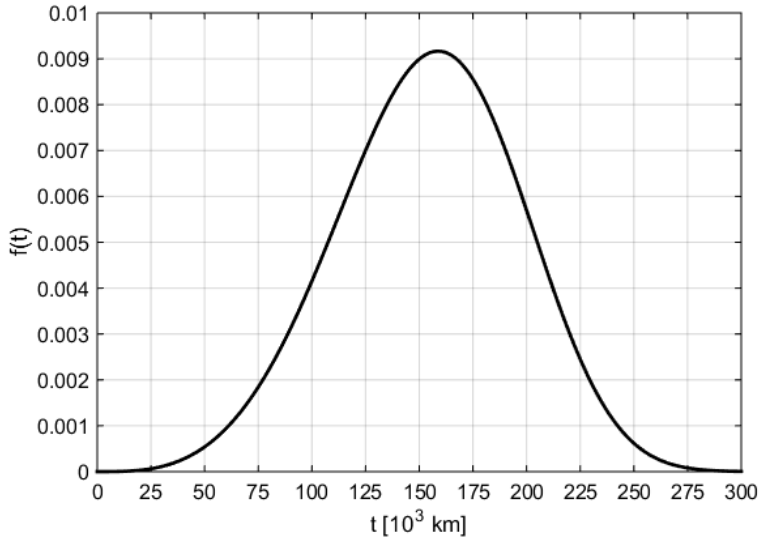
- w przypadku uszkodzenia elementu, następuje jego odnowa, bez odnowy elementów nieuszkodzonych oraz obliczenie chwili najbliższej wymaganej odnowy profilaktycznej wszystkich elementów,
- wykonywanie odnowy prewencyjnej wszystkich elementów w chwili wyznaczanej każdorazowo po awarii dowolnego elementu (jak wyżej).

W tym przypadku jedyną różnicą w odniesieniu do danych wejściowych jest brak konieczności określenia wartości granicznej czasu t_{i_gr}.

4. Wyniki obliczeń

Obliczenia wykonano na podstawie danych inspirowanych doświadczeniami z praktyki eksploatacyjnej monoblokowych zestawów kołowych elektrycznych zespołów trakcyjnych (EZT) składających się z kilku wagonów. Przyjęto, że rozkład ich czasu pracy do uszkodzenia jest zgodny z rozkładem Weibulla o parametrze kształtu υ = 4,1 i parametrze skali β = 170000 [km]. Funkcję gęstości prawdopodobieństwa przebiegu do uszkodzenia dla tego rozkładu zapisano zależnością (3) i przedstawiono na rysunku 5.

$$\mathrm{f}(t) = \upsilon \left(\frac{1}{\beta}\right)^{\upsilon} t^{\upsilon-1} exp\left(-\left(\frac{t}{\beta}\right)^{\upsilon}\right) \tag{3}$$



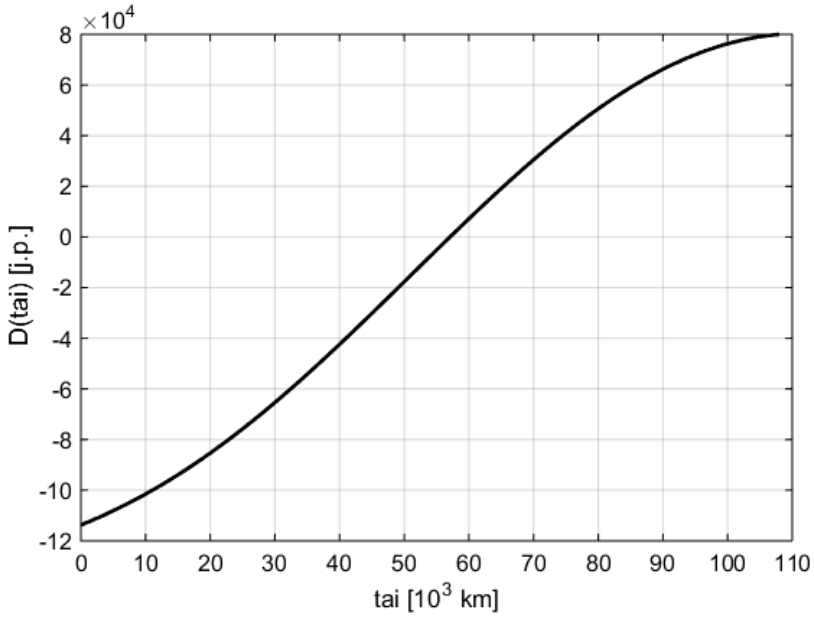
Rys. 5. Przyjęta funkcja gęstości prawdopodobieństwa czasu pracy do uszkodzenia analizowanych zestawów kołowych

Rozważane są tylko uszkodzenia wynikające ze zużycia kół pojazdów w czasie eksploatacji, polegające na przekroczeniu granicznych wartości wymiarów profilu koła określonych w dokumentacji systemu utrzymania pojazdu (DSU). Realizowany u przewoźnika proces eksploatacji zestawów kołowych obejmuje przeglądy i odnowy wykonywane zgodnie z

DSU. Przeglądy przeprowadzane są w z góry ustalonych chwilach – po określonym przebiegu i czasie.

Uwzględniając koszty: odnowy prewencyjnej – na poziomie k_o = 80000 [j.p. – jednostek pieniężnych] i odnowy poawaryjnej – na poziomie k_a = 250000 [j.p.], wyznaczono optymalny okres odnowy prewencyjnej analizowanych zestawów kołowych. Wykorzystano w tym celu modele decyzyjno-losowe [13, 18] i uzyskano wartość czasu odnowy (wyrażoną w kilometrach przebiegu) t^* = 108000 [km].

Wykorzystując te dane wykreślono następnie krzywą obrazującą przebieg funkcji $D(tai)$ według równania (2), przedstawioną na rysunku 6.



Rys. 6. Przebieg funkcji $D(tai)$ dla analizowanego zestawu kołowego

Wynika z niej, że graniczna wartość przebiegu t_{i_gr} dla zestawu kołowego jest równa 58000 [km]. Jeśli koło ulegnie uszkodzeniu przed osiągnięciem tego przebiegu, miałoby być odnawiane po uszkodzeniu razem z pozostałymi kołami wagonu zespołu trakcyjnego i ponownie wraz ze wszystkimi kołami całego zespołu – również przy najbliższej wspólnej odnowie. Jeśli uszkodzi się po tym czasie jest odnawiane po uszkodzeniu, ale przy najbliższej wspólnej odnowie kół całego zespołu już nie – jego czas pracy jest wydłużony do kolejnej odnowy po najbliższej wspólnej odnowie.

Wykorzystując opracowane modele symulacyjne przeprowadzono obliczenia przyjmując do wykonania symulacji dodatkowo następujące wartości parametrów: T_H = 1000000 [km], dt = 5 [km], liczba powtórzeń symulacji n = 10000.

Wybrane wyniki uzyskane po przeprowadzeniu symulacji zestawiono w tabeli 1. Za najistotniejszy wynik symulacji uznano koszty odnowy. Znajduje to uzasadnienie w praktyce eksploatacyjnej, w której często aspekt ekonomiczny eksploatacji stanowi podstawę do oceny wykorzystywanego układu i przyjętej strategii odnowy [5]. Uzyskiwane wyniki stanowią więc istotne przesłanki do podejmowania decyzji w trakcie kierowania procesem eksploatacji.

W tabeli 1 przedstawiono średnie koszty odnów prewencyjnych, poawaryjnych i ich wartość sumaryczną – uzyskane przy stosowaniu obydwu strategii odnawiania rozważanych zestawów kołowych w przyjętym horyzoncie czasowym symulacji.

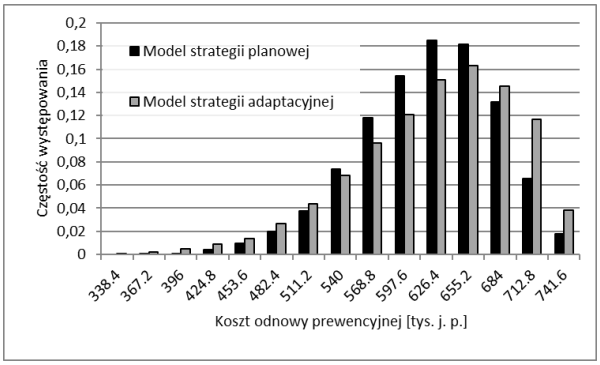
Tab. 1. Wyniki symulacji

Rodzaj wartości kosztów odnowy	Model strategii planowej	Model strategii adaptacyjnej
Średni koszt odnów prewencyjnych [j.p.]	614480	618962
Średni koszt odnów poawaryjnych [j.p.]	1093025	1069500
Średni koszt całkowity odnów [j.p.]	1707505	1688462

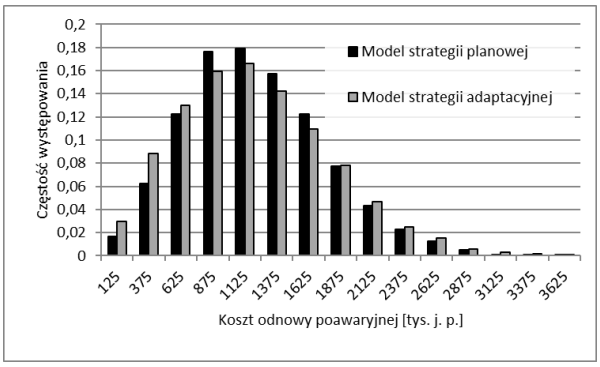
Analizując uzyskane wyniki, można zauważyć, że w rozpatrywanym przypadku mniejsze koszty łączne otrzymano przy stosowaniu strategii adaptacyjnej odnawiania, co sugeruje, iż jej stosowanie może być bardziej efektywne ekonomicznie w przypadku rzeczywistej eksploatacji modelowanego obiektu technicznego.

Istotną zaletą wykorzystania symulacji komputerowej do analizy pracy układu według przedstawianych modeli jest możliwość uzyskania rozkładów kosztów odnowy, które zestawiono na rysunku 7.

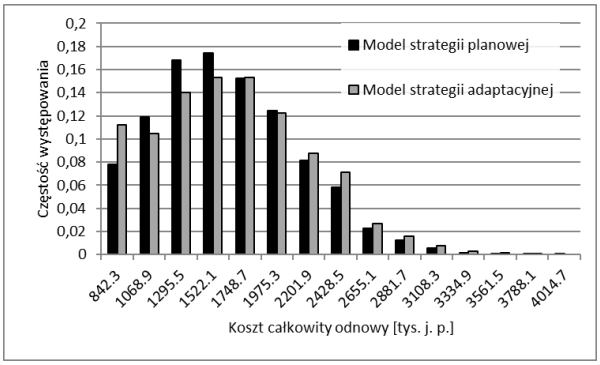
a)



b)



c)



Rys. 7. Rozkłady kosztów odnów prewencyjnych, poawaryjnych i całkowitych dla analizowanych strategii

Prowadzona symulacja daje zatem możliwość oceny prawdopodobieństwa, z jakim koszt odnowy układu nie przekroczy założonego poziomu.

5. Podsumowanie

Ekonomiczne wskaźniki procesu eksploatacji są jednymi z najważniejszych wskaźników wykorzystania pojazdów w systemach transportowych, bowiem to od ekonomiki eksploatacji w dużej mierze zależy poprawne funkcjonowanie całego przedsiębiorstwa. Zaproponowane w pracy metody w znaczący sposób mogą wpłynąć na racjonalne organizowane przeglądów i odnowy profilaktycznej. Wybór strategii obsługiwania powinien zatem uwzględniać kryterium efektywności ekonomicznej. Pomimo, że dla analizowanych strategiach występujące różnice ich stosowania nie są duże, to w praktyce prowadzą one do poważnych konsekwencji ekonomicznych ze względu na znaczące koszty działań obsługowych.

Opracowane i przedstawione w niniejszej pracy metody planowania odnowy prewencyjnej w systemach eksploatacji pojazdów stanowią przydatne narzędzie wspomagania decyzji i prowadzenia racjonalnej eksploatacji obiektów technicznych. Każdy z przedstawionych modeli jest przydatny do zastosowania w różnych warunkach eksploatacji oraz dla złożonych i zróżnicowanych konstrukcyjnie obiektów technicznych.

Przedstawiony w artykule adaptacyjny model wyznaczania strategii odnów prewencyjnych stanowi przyczynek do rozwoju metod przeprowadzania remontów i napraw złożonych obiektów technicznych.

Finansowanie badań

Praca została sfinansowana przez Akademię Górniczo-Hutniczą, Wydział Inżynierii Mechanicznej i Robotyki, program badawczy nr 16.16.130.942.

Literatura

1. Andrzejczak K, Młyńczak M, Selech J. Assessment model of operational effectiveness related to newly operated public means of transport. Proceedings of the 27th European Safety and Reliability Conference ESREL 2017: 3455–3461.
2. Badia F G, Berrade M D, Campos C A. Optimal inspection and preventive maintenance of units with revealed and unrevealed failures. Reliability Engineering and System Safety 2002; 78: 157–163.
3. Berrade M D, Scarf P A, Cavalcante C A V, Dwight R A. Imperfect inspection and replacement of a system with a defective state: A cost and reliability analysis. Reliability Engineering and System Safety 2013; 120: 80–87.
4. Bradley E. Reliability engineering a Life Cycle Approach. Boca Raton: CRC Press Taylor & Francis Group, 2017.
5. Faulin J, Juan Perez A A, Martorell Alsina S S, Ramirez-Marquez J E (Eds.). Simulation Methods for Reliability and Availability of Complex Systems. London, New York: Springer, 2010.

6. Macchi M, Garetti M, Centrone D, et al. Maintenance management of railway infrastructures based on reliability analysis. Reliability Engineering & System Safety 2012; 104: 71–83.
7. Młynarski S, Pilch R, Smolnik M, Szybka J, Wiązania G. A concept of reliability assessment simulation model using systems structural decomposition. Journal of KONBiN 2018; 46: 51–74.
8. Młynarski S, Pilch R, Smolnik M, Szybka J, Wiązania G. Formation of koon Systems Reliability Estimated with Analytical and Simulation Calculation Methods. Journal of KONBiN 2017; 42: 255–272.
9. Młyńczak M. Failure models of mechanical objects. Zagadnienia Eksploatacji Maszyn 2010; 45: 29–43.
10. Nachlas J A. Reliability engineering. Probabilistic models and maintenance methods. Boca Raton: CRC Press Taylor & Francis Group, 2017.
11. O'Connor P. Practical reliability engineering. Chichester: John Wiley & Sons Ltd., 2012.
12. Peng W, Huang H Z, Zhang X, Liu Y, Li Y. Reliability based optimal preventive maintenance policy of series-parallel systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2009; 2: 4–7.
13. Pilch R. Determination of preventive maintenance time for milling assemblies used in coal mills. Journal of Machine Construction and Maintenance – Problemy Eksploatacji 2017; 1: 81–86.
14. Saranga H, Kumar U D. Optimization of aircraft maintenance/support infrastructure using genetic algorithms level of repair analysis. Annals of Operations Research 2006; 143: 91–106.
15. Serkan E, Yilser D. Reliability and optimal replacement policy for a k-out-of-n system subject to shocks. Reliability Engineering & System Safety 2019; 188: 393–397.
16. Song H, Schnieder E. Modeling of railway system maintenance and availability by means of colored Petri nets. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 2: 236–243.
17. Sowa A. Formal models of generating checkup sets for the technical condition evaluation of compound objects. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16(1): 150–157.
18. Smolnik M. Projektowanie procesu obsługiwania obiektów technicznych na przykładzie wybranych wagonów tramwajowych [PhD thesis]. Kraków: AGH w Krakowie, 2018.
19. Świderski A, Jóźwiak A, Jachimowski R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 2: 292–299.
20. Werbińska-Wojciechowska S. Preventive Maintenance Models for Technical Systems. In: Technical System Maintenance. Springer Series in Reliability Engineering. Springer, Cham, 2019.
21. Zhao Y X. On preventive maintenance policy of a critical reliability level for system subject to degradation. Reliability Engineering and System Safety 2003; 79: 301–308.

dr inż. Łukasz Sobaszek
dr inż. Arkadiusz Gola
prof. dr hab. inż. Antoni Świć
Katedra Informatyzacji i Robotyzacji Produkcji
Wydział Mechaniczny, Politechnika Lubelska
ul. Nadbystrzycka 36, 20-816 Lublin, Polska
e-mail: l.sobaszek@pollub.pl, a.gola@pollub.pl, a.swic@pollub.pl

ALGORYTM WSPARCIA STRATEGII TBM W WIELOMASZYNOWYCH SYSTEMACH WYTWÓRCZYCH

TIME-BASED PREDICTION OF MACHINE FAILURES IN MULTI-MACHINES MANUFACTURING SYSTEMS

Streszczenie: *Realizacja procesów produkcyjnych w rzeczywistych systemach wytwórczych wiąże się z występowaniem wielu zakłóceń, do których zalicza się głównie awarie maszyn technologicznych. W związku z tym obserwowany jest rozwój różnorodnych strategii utrzymania ruchu. Coraz większy nacisk kładziony jest na efektywne działania prewencyjne, do których zalicza się także działania określone w czasie (ang. Time-Based Maintenance – TBM). W niniejszej publikacji zaprezentowano algorytm predykcji awarii maszyn w wielomaszynowych systemach wytwórczych wspierający prewencyjne utrzymanie ruchu. Na wstępie omówiono zagadnienia związane z typowymi strategiami stosowanymi w obszarze UR. Ponadto omówiono tematykę predykcji awarii, zwracając uwagę na ujęcie czasu pracy maszyny jako czasu trwania, a także kluczową rolę wykorzystania danych historycznych dotyczących awarii maszyn. Następnie zaprezentowano proponowany algorytm predykcji wspierający działania określone w czasie. Prezentowane prace zakończono dwuetapową weryfikacją proponowanej metody, która potwierdziła jej skuteczność oraz zasadność wykorzystania.*

słowa kluczowe: *system produkcyjny, utrzymanie ruchu, niezawodność, awarie maszyn, predykcja, Time-Based Maintenance*

1. Wstęp

Rzeczywistość produkcyjna związana jest z występowaniem wielu zakłóceń, które w negatywny sposób mogą wpływać na realizowane procesy, powodując ich dezorganizację [14]. Do kluczowych czynników niepewności należy zaliczyć występowanie awarii maszyn technologicznych. Z praktycznego punktu widzenia, określenie czasu wystąpienia awarii jest niezwykle ważnym zagadnieniem. Znajomość czasów występowania usterek pomaga w podejmowaniu przemyślanych działań prewencyjnych – należy je bowiem planować tak, aby nie kolidowały one z bieżącą realizacją procesu produkcyjnego. Predykcja czasów awarii znajduje zastosowanie w obszarze strategii Time-Based Maintenance (TBM), gdzie coraz częściej proponowane są narzędzia informatyczne wspierające tą strategię [5, 16, 37]. Istotne jest jednak, aby proponowane rozwiązania wykorzystywały efektywne algorytmy predykcji wykorzystujące rzetelne dane historyczne na podstawie których możliwa jest wiarygodna analiza występowania usterek, a w konsekwencji planowanie odpowiednich działań serwisowych [6, 13, 40].

W literaturze przedmiotu znaleźć można wiele opracowań podejmujących problematykę predykcji zakłóceń procesu produkcyjnego. W ogólnym ujęciu, w obszarze utrzymania ruchu prowadzi się badania w kierunku opracowywania efektywnych metod zapobiegania występowaniu usterek, a także niwelowania ich wpływu [3, 33]. Planowanie działań prewencyjnych przyjmuje zazwyczaj formę podejmowania działań określonych w czasie (ang. *Time-Based Maintenance*) [13, 25], a także działań opartych na monitorowaniu warunków pracy maszyn (ang. *Conditioned-Based Maintenance*) [1, 30]. Wyraźnym trendem jest także opracowywanie scenariuszy oraz strategii eksploatacyjnych [26, 27, 34, 35, 39].

Proponowane w literaturze metody predykcji awarii można podzielić na klika grup, wśród których wyróżnia się:

- metody bazujące na wykorzystaniu rozkładów prawdopodobieństwa,
- metody wykorzystujące typowe wskaźniki efektywności,
- alternatywne metody predykcji awarii,
- metody bazujące na wykorzystaniu danych rzeczywistych.

Zdecydowana większość proponowanych w literaturze rozwiązań opiera się na analizach rozkładów prawdopodobieństwa [8, 15, 24, 2]. Rozpatrywane są zarówno typowe rozkłady, jak i ich kombinacje. Zastosowanie znajdują różnorodne rozwiązania – od wykorzystania rozkładu jednostajnego [17, 2], poprzez bazowanie na rozkładzie normalnym [8], aż po predykcję na podstawie rozkładu wykładniczego [24, 30]. Analiza rozkładów pozwala głównie zdefiniować czas wystąpienia awarii. W literaturze proponowane są także rozwiązania bazujące na kombinacjach typowych rozkładów. Przykładowo w pracy [15] autorzy do opisu problemu występowania awarii proponują użycie kombinacji rozkładów: normalnego, trójkątnego i wykładniczego. Większość z proponowanych rozwiązań ma jednak wciąż charakter rozważań teoretycznych. Pomijany jest przez to kluczowy aspekt wykorzystania historycznych danych dotyczących awaryjności wykorzystywanych maszyn. Ponadto badacze nie argumentują w dostateczny sposób doboru rozpatrywanych rozkładów.

Kolejnym, widocznym w literaturze trendem jest wykorzystywanie dla celów predykcji typowych wskaźników efektywności, stosowanych w obszarze utrzymania ruchu takich jak:

- średni czas do wystąpienia uszkodzenia MTTF (ang. *Mean Time To Failure*),
- średni czas bezawaryjnej pracy MTBF (ang. *Mean Time Between Failures*),
- średni czas naprawy MTTR (ang. *Mean Time To Repair*).

Wykorzystanie wspomnianych wskaźników prezentowane jest wielu w pracach [9, 12, 21, 20]. Są one zazwyczaj stosowane bezpośrednio, jednak zdarza się, iż służą w pośredni sposób do oszacowania parametrów rozkładu Weibulla [21]. W ramach realizowanych badań autorzy analizują odpowiednio opracowane scenariusze, zaś wartości wskaźników przyjmują wartości z uprzednio zdefiniowanych przedziałów – dobierane są tak, aby zdarzenia występowały często, czasami lub rzadko, a następnie analizowane są konsekwencje występowania awarii [12]. Niekiedy wykorzystanie wskaźników efektywności wspierane jest zastosowaniem odpowiednich metod statystycznych [30]. Metody mające na celu wykorzystanie typowych wskaźników efektywności stosowanych w obszarze utrzymania ruchu wydają się być zasadne, gdyż parametry te dostarczają wiele informacji nt. wykorzystywanych maszyn technologicznych. Wciąż jednak pozyskiwanie i wykorzystanie rozpatrywanych wielkości traktowane jest dość teoretycznie. W publikacjach, w których proponowane są metody wykorzystujące typowe wskaźniki efektywności, brak jest weryfikacji proponowanych rozwiązań z zastosowaniem rzeczywistych danych dotyczących awaryjności maszyn [9, 20].

W obszarze alternatywnych metod predykcji awarii znaleźć można także podejścia, które znacznie różnią się od typowych rozwiązań. Warto wyróżnić tutaj rozwiązania w których: awarie zostają skumulowane do jednej i opisywane są za pomocą parametru MTTR oraz stopnia awarii MBL (ang. *Machine Breakdown Level*) [18], awaryjność określana jest na podstawie analizy rozkładów czasów obciążenia maszyn [31], do predykcji awarii maszyn wykorzystane zostają sztuczne sieci neuronowe [4], czy dobrze znane modele szeregów czasowych [38]. Podczas weryfikacji proponowanych rozwiązań wykorzystywane są jednak zazwyczaj dane testowe, często przyjęte przez autorów na podstawie opracowanych założeń upraszczających.

W publikacjach [8, 19] autorzy zwracają uwagę, iż istotną kwestią jest bazowanie na zbiorach rzeczywistych danych dotyczących procesu. Jest to sugestia, która stanowi nowe podejście w procesie predykcji awarii. Determinuje ona potrzebę opracowywania metod zorientowanych na praktyczne wykorzystanie historycznych danych dotyczących awaryjności

maszyn technologicznych. W literaturze można spotkać takie rozwiązania, jednak ich liczba jest wciąż bardzo mała [33]. Jest to wyraźny trend w obszarze predykcji awarii.

Niestety mimo wielu proponowanych przez badaczy metod, brak jest rozwiązań zorientowanych na praktyczne wykorzystanie historycznych danych dotyczących awaryjności maszyn technologicznych. Ponadto w praktyce produkcyjnej proponuje się niejednokrotnie wdrażanie rozbudowanych i kosztownych systemów monitorowania, podczas gdy w zakresie działań związanych ze strategią TBM zastosowanie zależć mogą dane posiadane przez wszystkie działy utrzymania ruchu. Dlatego też, w niniejszej pracy zaproponowany został algorytm predykcji awarii maszyn w wielomaszynowych systemach wytwórczych, który poprzez wnikliwą analizę rzeczywistych danych produkcyjnych pozwala na wnioskowanie o przyszłych czasach występowania usterek oraz podejmowanie skutecznych działań prewencyjnych. Jest on rozwiązaniem alternatywnym w stosunku do metod już istniejących, które zostały scharakteryzowane powyżej. Proponowane rozwiązanie pozwala na wykorzystanie danych posiadanych przez służby utrzymania ruchu do osiągnięcia zamierzonego celu w postaci identyfikacji możliwego momentu powstania awarii. Nowatorskim elementem proponowanej metody jest wykorzystanie elementów teorii analizy przeżycia w obszarze analizy awaryjności maszyn technologicznych, umożliwiającej wnioskowanie na podstawie danych historycznych.

2. Predykcja awarii z wykorzystaniem elementów analizy czasu trwania

2.1. Czas pracy maszyny jako czas trwania

Problem predykcji awarii polega w znacznym uproszczeniu na wyznaczeniu czasu w którym może wystąpić potencjalna usterka danej maszyny technologicznej wraz z prawdopodobieństwem jej wystąpienia. Do tego celu wykorzystać można elementy Analizy Przeżycia, nazywanej także Analizą Czasu Trwania [11, 23].

Ważnym elementem wykorzystania Analizy Czasu Trwania jest dokładne ustalenie istoty badanego procesu, który powinien spełniać następujące warunki [11]:

1. Zmiany dotyczące jednostki objętej badaniem muszą dokonywać się pomiędzy stanami dyskretnymi.
2. Zmiany stanów mogą występować w dowolnym momencie i nie są z góry ograniczone ustalonymi punktami w czasie.
3. Zmiany mogą być odwracalne lub nieodwracalne (w zależności od formy procesu).
4. Zmiany są ustalone z góry przez stan, w jakim znajduje się proces.
5. Istnieją czynniki wpływające na proces, a analiza pozwala na ich wykrycie.

Analizując powyższe na tle procesu występowania awarii maszyn technologicznych, należy stwierdzić, iż proces ten spełnia te wymagania. Awaria maszyny może wystąpić w dowolnej chwili i jest zmianą zachodzącą pomiędzy dwoma stanami – sprawna i uszkodzona. Ponadto uszkodzenie maszyny jest zmianą odwracalną – naprawa powoduje powrót do pierwotnego stanu, który jest definiowany przez stan, w jakim znajduje się urządzenie. Istnieje także szereg czynników, które mogą wpływać na analizowany proces i mogą być identyfikowane dzięki wykorzystaniu Analizy Czasu Trwania [36]. Czas trwania w przypadku maszyny należy rozumieć jako czas jej sprawnej pracy, na podstawie którego możliwe jest również wnioskowanie o czasie wystąpienia awarii. W konsekwencji atutem tej techniki jest możliwość wyznaczania wzorców występowania awarii (charakterystyk występowania awarii w czasie) zwłaszcza wtedy, gdy posiadane dane historyczne nie pozwalają na zastosowanie typowych technik wnioskowania [33].

Niech zatem T będzie nieujemną zmienną losową reprezentującą czas wystąpienia awarii maszyny technologicznej, która przyjmuje wartości z przedziału $(0; \infty)$. Ponadto $f(t)$

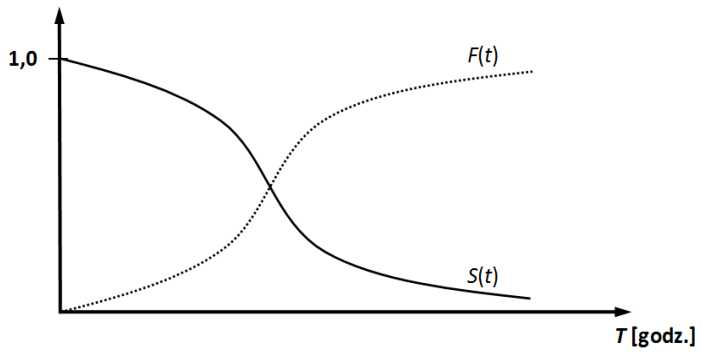
jest funkcją gęstości prawdopodobieństwa, gdzie $t > 0$ oraz $F(t)$ jest dystrybuantą zmiennej losowej T – funkcją niemalejącą określającą, iż obiekt doświadczy zdarzenia w czasie (0; t]:

$$F(t) = P(T < t). \tag{1}$$

Bazując na dystrybuancie $F(t)$ zdefiniować można funkcję przeżycia $S(t)$:

$$S(t) = 1 - F(t) = P(T \geq t) = \int_t^{\infty} f(s)ds, \tag{2}$$

pozwalającą określić prawdopodobieństwo poprawnej pracy maszyny, aż do chwili t. Pozwala ona zatem określić także prawdopodobieństwo, iż awaria nie wystąpi do tego czasu. Wyznaczona funkcja jest doskonałym sposobem określania wzorców poprawnej pracy maszyny, a w konsekwencji także występowania jej awarii. Funkcja przeżycia oraz dystrybuanta zostały przedstawione na rysunku 1.



Rys. 1. Dystrybuanta F(t) oraz funkcja przeżycia S(t)

W celu wyznaczenia poszczególnych funkcji zaprezentowanych powyżej, wykorzystać należy odpowiednie dane historyczne dotyczące awaryjności maszyny technologicznej. Ich analiza dostarcza bowiem wielu informacji, które mogą zostać wykorzystane w dalszym procesie predykcji.

2.2. Wykorzystanie danych historycznych

Do wyznaczenia charakterystyk występowania awarii, niezbędne jest zdefiniowanie odpowiedniego źródła danych. Takie informacje gromadzone są zazwyczaj przez działy utrzymania ruchu przedsiębiorstw produkcyjnych [3, 10]. Dokonując analizy rozwiązań dotyczących zapisu wiedzy z zakresu historii konserwacji i napraw maszyn technologicznych stosowanych w przedsiębiorstwach produkcyjnych należy stwierdzić, iż najczęściej stosowanymi rozwiązaniami są:

- prowadzenie papierowej dokumentacji serwisowej – najczęściej w formie Kart Obsługi oraz Książek Serwisowych,
- wykorzystanie w procesie gromadzenia informacji serwisowych oprogramowania komputerowego wraz z dedykowanymi arkuszami danych (rys. 2),
- gromadzenie danych bezpośrednio z maszyn technologicznych, z wykorzystaniem systemów klasy SCADA (ang. *Supervisory Control And Data Acquisition*) oraz MES (ang. *Manufacturing Execution Systems*).

Machine	Shift	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
		Śr	Cz	Pt	So	Nd	Pn	Wt	Śr	Cz	Pt	So	Nd	Pn	Wt	Śr	Cz	Pt	So	Nd	Pn	Wt	Śr	Cz	Pt	So	Nd	Pn	Wt	Śr	Cz
1	I	0,67	8	0,08						8						0,3	8	0,33						8	1				2,67	0,67	8
	II						0,5			3,5						1,17	0,92	2,33												2	3,8
	III		2,33	0,17												0,42								0,33						1,08	0,3
2	I		8							4				0,33			8				8			8							8
	II	2,5	1,08											0,17																	
	III		1																				1,3					1,42			
3	I	8						8	1,3		0,3				8							8						0,5		8	
	II	0,33	1	0,5				0,33		1,17	0,33			1		1,17													1		
	III			0,17						1,5	0,17																	2		0,8	
4	I			1,17					6		0,58					8		0,5					8	1,83	1,92				8		
	II							0,3			1,17	0,3		1,67		1							0,33	1							
	III	0,8					2,33					0,33			0,5	0,58								1,8	1			0,67			

Rys. 2. Przykład danych serwisowych zapisanych z wykorzystaniem komputerowego arkusza danych

Wszystkie z przedstawionych powyżej metod gromadzenia danych łączy wspólna cecha – każda z nich dostarcza danych, które odpowiednio przetworzone mogą zostać wykorzystane w procesie predykcji awarii maszyn z zastosowaniem elementów Analizy Przeżycia.

Podstawową informację zawartą we wspomnianej dokumentacji są historyczne czasy występowania usterek. Dla danej maszyny technologicznej M_j można zapisać je jako zbiór danych T_{Mj}:

$$T_{Mj} = \{t_1, t_2, ..., t_n\} \text{ [godz.]}, \tag{3}$$

gdzie: t_i – i-ty czas wystąpienia awarii.

Przykładowy zbiór historycznych danych dotyczący czasów awarii dla maszyny M_1 przyjmie postać:

$$T_{M1} = \{4, 8, 20, 16, 10, 28, 43, 15, 24, 2, ...\} \text{[godz.]}.$$

Wykorzystanie w procesie predykcji danych zawartych w odpowiednich zbiorach T_{Mj} pozwala na określenie potencjalnych czasów wystąpienia awarii danej maszyny, które zapisać można w zbiorze FT_{Mij}:

$$FT_{Mj} = \{ ft_{Mj1}, ft_{Mj2}, ..., ft_{Mjn} \}, \tag{4}$$

gdzie: ft_{Mji} – czas wystąpienia awarii maszyny j,
j – numer rozpatrywanej maszyny technologicznej.

Dla każdego czasu ft_{Mji} wyznaczone zostanie ponadto prawdopodobieństwo wystąpienia awarii zapisane w zbiorze P_{Mj}.

$$P_{Mj} = \{ p_{Mj1}, p_{Mj2}, ..., p_{Mjn} \}, \tag{5}$$

gdzie: p_{Mji} – wartość prawdopodobieństwa wystąpienia awarii maszyny j, przy czym:

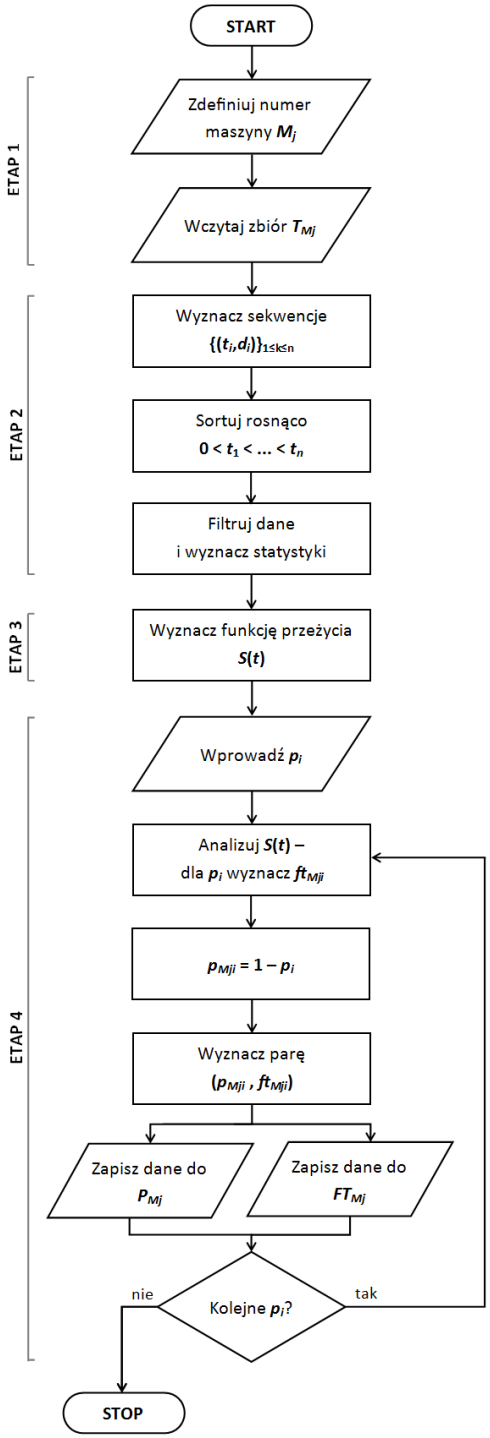
$$\bigwedge_{ft_{Mij} \neq 0} p_{Mij} \neq 0 .$$

Zatem rezultatem procesu predykcji będzie wyznaczenie par (p_{Mji}, ft_{Mji}) definiujących prawdopodobieństwo oraz czas wystąpienia awarii maszyny M_j.

2.3. Proponowany algorytm predykcji wspierający działania TBM

W celu predykcji szukanych wartości prawdopodobieństwa wystąpienia awarii oraz czasu usterki opracowano czteroetapowy algorytm umożliwiający analizę oraz odpowiednie wykorzystanie zgromadzonych danych dotyczących historii napraw (rys. 3).

Etap 1 proponowanego algorytmu polega na zdefiniowaniu maszyny dla której prowadzony będzie proces predykcji, a także zaimportowaniu danych historycznych zawartych w zbiorze T_{Mj}.



Rys. 3. Algorytm predykcji czasu wystąpienia awarii

W etapie 2 realizowany jest odpowiedni zapis zaimportowanych danych – czasy awarii wybranej maszyny technologicznej M_j zostają wówczas zapisane za pomocą odpowiedniej sekwencji:

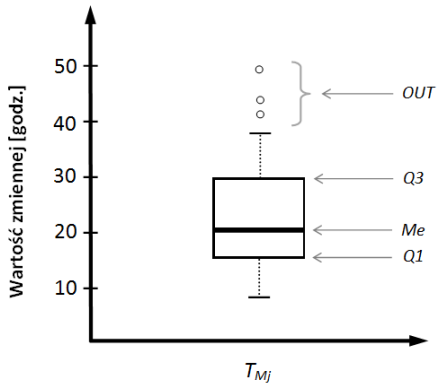
$$\{(t_i, d_i)\}_{1 \le k \le n}, t_i \in T_{Mj}, \tag{6}$$

gdzie: t_i – czas pomiędzy kolejnymi awariami,
d_i – liczba przypadków.

Ponadto na tym etapie realizowana jest operacja sortowania – sekwencje zostają uporządkowane według rosnących wartości $\{t_i\}_{1 \le k \le n}$:

$$0 < t_1 < t_2 < \ldots < t_n, \tag{7}$$

po czym następuje filtracja zgromadzonych danych i usunięcie obserwacji odstających (wartości nietypowych) (rys. 4). Następnie wyznaczone zostają podstawowe statystki dla zgromadzonych danych (wartość minimalna, maksymalna, średnia, rozstęp, kwartyle).



Rys. 4. Wykres pudełkowy dla przykładowych danych (Me – mediana, Q1 i Q3 – kwartyl 1 i 3, OUT – obserwacje odstające)

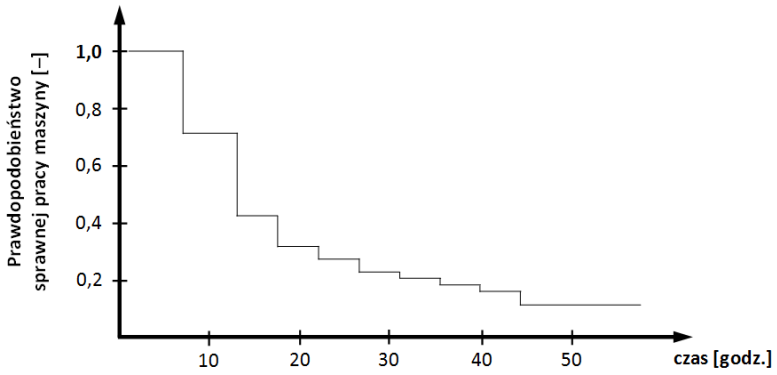
Etap 3 stanowi kluczowy element procesu wnioskowania, gdyż właśnie na tym etapie wyznaczana jest funkcja przeżycia charakteryzująca rozpatrywany proces awaryjności analizowanej maszyny. Uszeregowanie przypadków awarii według rosnących czasów ich wystąpienia, a także określenie liczby przypadków dla każdego z czasów pozwala na wyznaczenie funkcji przeżycia danego procesu. Wyznaczona postać funkcji przeżycia jest doskonałym sposobem określania wzorców występowania awarii – pozwala w przystępny sposób wyznaczyć charakterystyki awaryjności dla konkretnej maszyny technologicznej. Zastosowanie znajduje tu estymacja Kaplana-Meier'a – funkcja przeżycia wyznaczana jest wówczas na podstawie zależności:

$$\hat{S}(t) = \begin{cases} 1, & dla\ t < t_1, \\ \prod\limits_{t_i \le t} \dfrac{r_i - d_i}{r_i}, & dla\ t_1 < t, \end{cases} \tag{8}$$

gdzie: r_i – liczba wszystkich przypadków awarii określana jako:

$$r_i = \sum_{j=i}^{k} d_j\ . \tag{9}$$

Wówczas wyznaczona zostaje funkcja przeżycia na podstawie której (z określonym poziomem prawdopodobieństwa) określane są czasy sprawnej pracy maszyny (rys. 5).

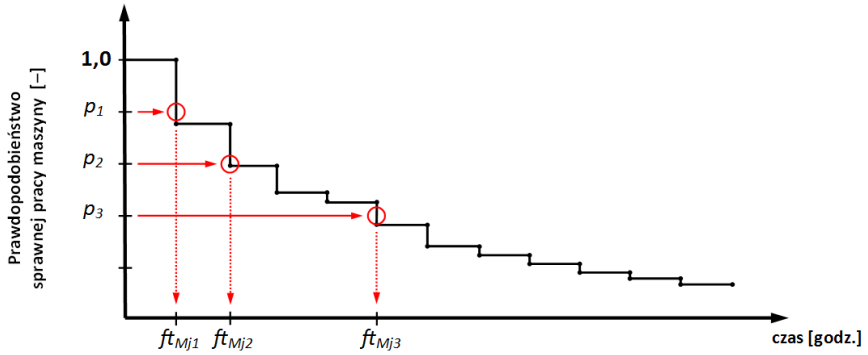


Rys. 5. Przykładowa funkcja przeżycia wyznaczona za pomocą estymacji Kaplana-Meier'a

W konsekwencji wyznaczona funkcja przeżycia wykorzystana zostaje na etapie 4, gdzie bazując na uzyskanych wynikach można wyznaczyć elementy poszukiwanych zbiorów:

– potencjalnych czasów wystąpienia awarii rozpatrywanej maszyny FT_{Mj},
– prawdopodobieństwa wystąpienia awarii rozpatrywanej maszyny P_{Mj}.

Idea procesu wnioskowania na podstawie wyznaczonej postaci funkcji przeżycia została przedstawiona na rys. 6. Dla odpowiednich poziomów prawdopodobieństwa p_i wyznaczone zostają prognozowane czasy wystąpienia awarii ft_{Mji}.



Rys. 6. Określenie czasu wystąpienia awarii na podstawie przyjętej wartości prawdopodobieństwa przeżycia

Z uwagi, iż na podstawie funkcji przeżycia określone są prawdopodobieństwa sprawnej pracy maszyny (p_i), zatem prawdopodobieństwo wystąpienia awarii p_{Mji} definiuje zależność:

$$p_{Mji} = 1 - p_i \,, \tag{10}$$

gdzie: p_{Mji} – prawdopodobieństwo doświadczenia usterki,
p_i – prawdopodobieństwo sprawnej pracy maszyny.

Określenie szukanych wartości prawdopodobieństwa wystąpienia awarii p_{Mji} umożliwia wyznaczenie szukanych wartości ft_{Mji}, a w konsekwencji wyznaczenie par (p_{Mji}, ft_{Mji}). Wyznaczone dane gromadzone są w zbiorach P_{Mji} oraz FT_{Mji}. Etap 4 ma charakter iteracyjny,

a zatem jest powtarzany w zależności od decyzji użytkownika co do ilości rozpatrywanych poziomów prawdopodobieństwa. Realizację algorytmu należy powtórzyć dla kolejnych maszyn technologicznych, dla których analizowana będzie ich awaryjność.

3. Weryfikacja eksperymentalna proponowanego algorytmu

3.1. Dane wykorzystane w procesie weryfikacji

Podstawą realizacji prezentowanej poniżej weryfikacji było pozyskanie i wykorzystanie odpowiednich danych, które dotyczyły zarówno procesów technologicznych, jak i awaryjności maszyn technologicznych. Przedstawione badania przeprowadzono w oparciu o rzeczywiste dane produkcyjne dotyczące realizacji 12 zadań produkcyjnych na 12 stanowiskach wytwórczych, zorganizowanych w postaci gniazd produkcyjnych. Wśród realizowanych procesów przeważa obróbka ubytkowa. Przykłady procesów technologicznych wybranych zadań produkcyjnych zostały przedstawione w tab. 1.

Tab. 1. Przykładowe procesy technologiczne zawarte w danych produkcyjnych

Nr wyrobu (zadania)	**Nr operacji**	**Stanowisko**	**Nazwa operacji**	tpz_{ij} [godz.]	to_{ij} [godz.]
1	10	Laser1	Cięcie blach	0,25	0,042
	20	Laser2	Cięcie rur i profili	0,20	0,017
	30	Prasa CNC	Gięcie krawędziowe	0,13	0,018
	40	Wiertarka	Wiercenie otworów	0,17	0,017
	50	Ślus.	Operacja ślusarska	0,08	0,017
	60	Spaw. – MIG	Spawanie MIG	0,13	0,092
3	10	Laser2	Cięcie laserem rur i profili	0,15	0,005
	20	Piła CNC	Cięcie na pile taśmowej	0,10	0,008
	30	Frezarka	Frezowanie	0,27	0,050
	40	Wiertarka	Wiercenie otworów	0,17	0,017
	50	Ślus.	Operacja ślusarska	0,08	0,033
	60	Spaw.– MIG	Spawanie MIG	0,13	0,033
	70	Tokarka	Toczenie	0,33	0,092
5	10	Laser1	Cięcie laserem blach	0,27	0,012
	20	Gilotyna	Cięcie na gilotynie	0,10	0,004
	30	Piła CNC	Cięcie na pile taśmowej	0,10	0,017
	40	Prasa CNC	Gięcie krawędziowe	0,17	0,025
	50	Wiertarka	Gwintowanie	0,13	0,100
	60	Ślus.	Operacja ślusarska	0,08	0,033
	70	Spaw.– TIG	Spawanie TIG	0,13	0,033
	80	Tokarka	Toczenie	0,33	0,108

Pozyskane dane rzeczywiste zostały wykorzystane w procesie weryfikacji w następującym zakresie:

- dane dotyczące awaryjności maszyn technologicznych posłużyły za dane wejściowe podczas weryfikacji opracowanego algorytmu,
- dane dotyczące realizowanych procesów wytwórczych zostały wykorzystane podczas badań symulacyjnych mających na celu ocenę skuteczności i zasadności wykorzystania proponowanego algorytmu w realnych warunkach produkcyjnych (z uwzględnieniem występowania awarii maszyn technologicznych).

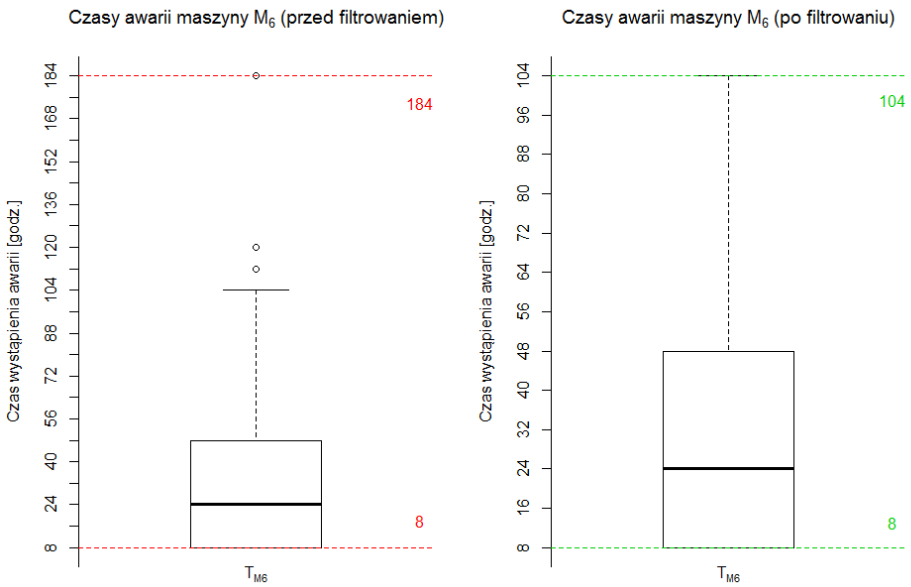
3.2. Predykcja czasu wystąpienia awarii

W celu weryfikacji proponowanego algorytmu opracowano w języku R odpowiedni skrypt umożliwiający jego realizację. Następnie z wykorzystaniem pozyskanych danych historycznych przeprowadzono proces wnioskowania o potencjalnych czasach awarii maszyn znajdujących się na poszczególnych stanowiskach. Dla poszczególnych maszyn przyjęto odpowiednie oznaczenia:

- Laser 1 – maszyna M_1,
- Laser 2 – maszyna M_2,
- Prasa CNC – maszyna M_3,
- Piła CNC – maszyna M_4,
- Stanowisko ślusarskie – maszyna M_5,
- Spawalnia – MIG – maszyna M_6,
- Spawalnia – TIG – maszyna M_7,
- Wiertarka – maszyna M_8,
- Frezarka – maszyna M_9,
- Tokarka – maszyna M_{10},
- Gilotyna – maszyna M_{11},
- Wykrawarka – maszyna M_{12}.

W dalszej części przedstawiono przykładową realizację procesu predykcji awarii dla maszyny M_6, w przypadku której dane historyczne liczyły 121 obserwacji

Przed rozpoczęciem procesu predykcji w skrypcie zawarto odpowiednie polecenia przygotowujące środowisko do pracy, po czym następuje zdefiniowanie numeru maszyny, oraz wczytanie danych zgromadzonych w pliku CSV. Import danych do zbioru T_{M6} przechowywanej w przestrzeni roboczej umożliwił realizację procesu sortowania obserwacji rosnąco, a także filtrację danych w której zastosowanie znalazły wygenerowane wykresy pudełkowe (rys. 7). Dodatkowo wyznaczone zostały podstawowe statystyki (rys. 8).



Rys. 7. Wykresy pudełkowe – przed oraz po filtrowaniu danych

```
 "Awaryjność maszyny M6 - zbiór TM6 - podstawowe statystyki:"
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
8.00    8.00   24.00   32.48   48.00  104.00
```

Rys. 8. Podstawowe statystki wygenerowane w ramach opracowanego skryptu

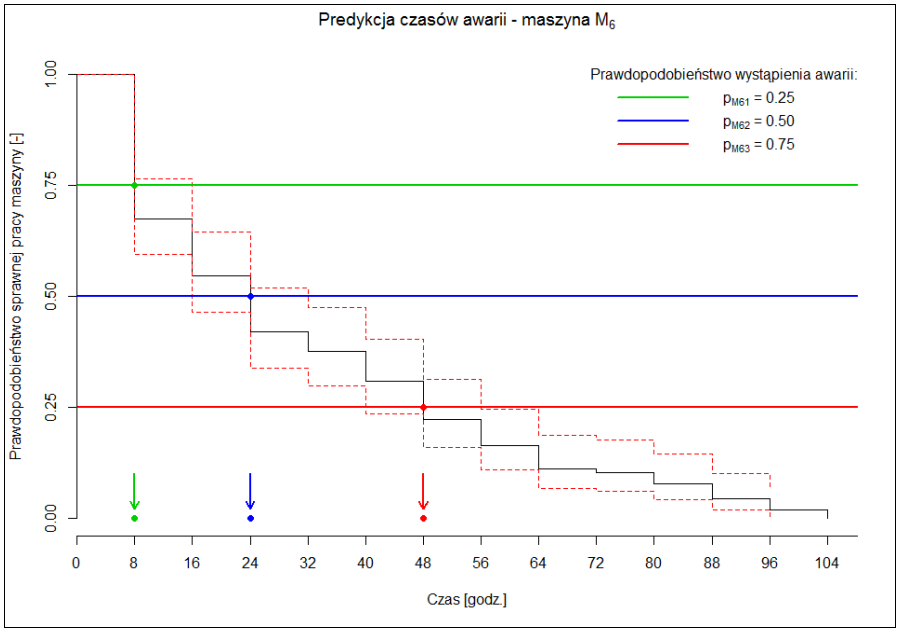
Kluczowym etapem algorytmu predykcji awarii jest wyznaczenie funkcji przeżycia $\hat{S}(t)$ z wykorzystaniem estymacji Kaplana-Meier'a. W opracowanym skrypcie była ona możliwa dzięki zastosowaniu biblioteki „SURVIVAL". Kolejne różnice krzywej schodkowej obliczane zostały w sposób automatyczny na podstawie utworzonych sekwencji obserwacji. Rezultatem było wyznaczenie funkcji przeżycia w postaci krzywej schodkowej z 95% przedziałem ufności.

Wyznaczenie przebiegu szukanej funkcji $\hat{S}(t)$, umożliwiło rozpoczęcie kolejnego etapu algorytmu jakim jest predykcja wartości czasu wystąpienia awarii rozpatrywanej maszyny dla zdefiniowanych poziomów prawdopodobieństwa (rys. 9). Ponieważ na wykresie odczytać można prawdopodobieństwo sprawnej pracy maszyny, dodatkowo generowana była legenda pomocnicza z objaśnieniami. W przypadku obliczeń dla danych maszyny M_6 (jak i pozostałych maszyn) rozpatrywano następujące poziomy prawdopodobieństwa:

$$p_1 = 0{,}75; \qquad p_2 = 0{,}50; \qquad p_3 = 0{,}25.$$

Wartości rozpatrywanych poziomów zostały zaś dobrane tak, aby wyznaczały: niski, średni oraz wysoki stopień ryzyka wystąpienia usterki rozpatrywanej maszyny. Stąd też:

$$p_{M61} = 1 - p_1 = 0{,}25; \quad p_{M62} = 1 - p_2 = 0{,}50; \quad p_{M63} = 1 - p_3 = 0{,}75.$$



Rys. 9. Predykcja awarii na podstawie wyznaczonej funkcji przeżycia

W ten sposób wyznaczone zostały szukane wartości prawdopodobieństw wystąpienia awarii oraz czasów potencjalnych usterek, które zapisać można jako pary:

$$(p_{M61}, ft_{M61}) = (0{,}25,\ 8 \text{ godz.}),$$
$$(p_{M62}, ft_{M62}) = (0{,}50,\ 24 \text{ godz.}),$$
$$(p_{M63}, ft_{M63}) = (0{,}75,\ 48 \text{ godz.}).$$

W konsekwencji wyznaczono zbiory P_{M61} = {0,25, 0,50, 0,75} oraz FT_{M61} = {8, 24, 48} [godz.].

Proponowany algorytm wykorzystano w takim samym zakresie dla pozostałych maszyn technologicznych. Jedynie z uwagi na charakter stanowiska ślusarskiego (M_5) proces predykcji w tym przypadku nie był realizowany. Uzyskane czasy wystąpienia awarii zostały zestawione w tab. 2.

Tab. 2. Czasy awarii maszyn technologicznych uzyskane w wyniku predykcji

	Czas wystąpienia awarii [godz.]		
Maszyna	p_{Mj1} **= 0,25**	p_{Mj2} **= 0,50**	p_{Mj3} **= 0,75**
M_1	8	16	40
M_2	8	24	32
M_3	8	16	24
M_4	8	24	104
M_5	–	–	–
M_6	8	24	48
M_7	8	16	40
M_8	8	24	48
M_9	8	16	40
M_{10}	8	24	40
M_{11}	8	16	40
M_{12}	8	16	32

Uzyskane rezultaty realizacji algorytmu zostały wykorzystane w dalszej części procesu weryfikacji polegającej na symulacji produkcji w warunkach występowania awarii maszyn technologicznych.

3.3. Symulacja produkcji z warunkach niepewności

W celu weryfikacji proponowanego algorytmu w realnych warunkach produkcyjnych, związanych z występowaniem niepewności dotyczącej awaryjności maszyn technologicznych, przeprowadzono dwuetapowy eksperyment w którym:

1. Dla rzeczywistych danych produkcyjnych zbudowano harmonogramy produkcji (harmonogramy nominalne), a następnie opracowano ich odpowiedniki z zaimplementowanymi buforami serwisowymi (harmonogramy odporne) w miejscach, na jakie wskazywały rezultaty zastosowania opracowanego algorytmu.
2. Przeprowadzono symulację produkcji zgodnie z opracowanymi harmonogramami, a następnie zbadano, które z harmonogramów wskazywały bliższy termin zakończenia produkcji w warunkach, gdy wystąpić mogą awarie maszyn.

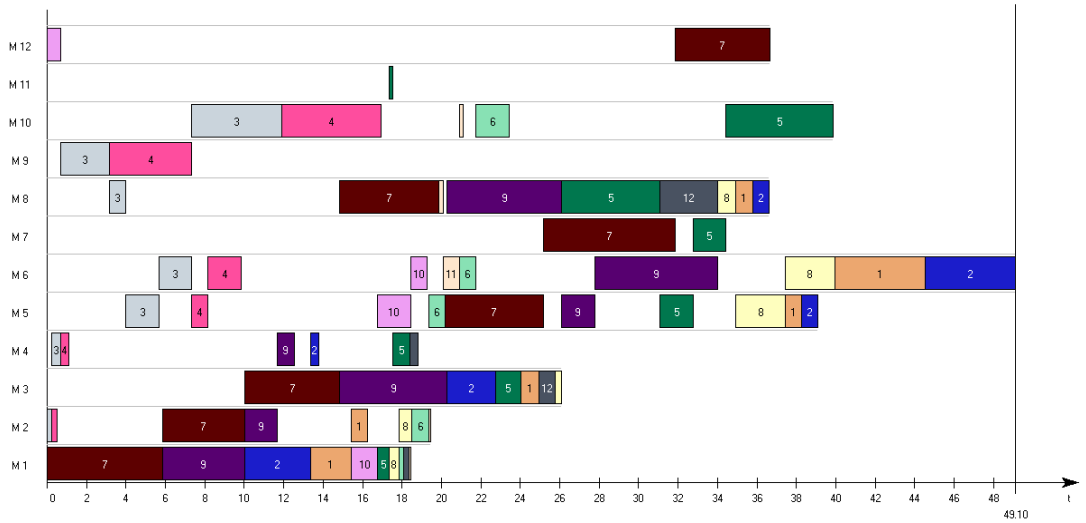
3.3.1. Opracowanie harmonogramów produkcji

W celu weryfikacji metody dla różnego uszeregowania zadań na poszczególnych stanowiskach produkcyjnych w procesie budowania harmonogramów wykorzystano 4 popularne reguły priorytetów:

1. Regułę FCFS (*First Come First Service – Pierwsza Przybyła Pierwsza Obsłużona*).
2. Regułę EDD (*Earliest Due Date – Najwcześniejszy Termin Dyrektywny*).
3. Regułę SPT (*Shortest Processing Time – Najkrótszy Czas Operacji*).
4. Regułę LPT (*Longest Processing Time – Najdłuższy Czas Operacji*).

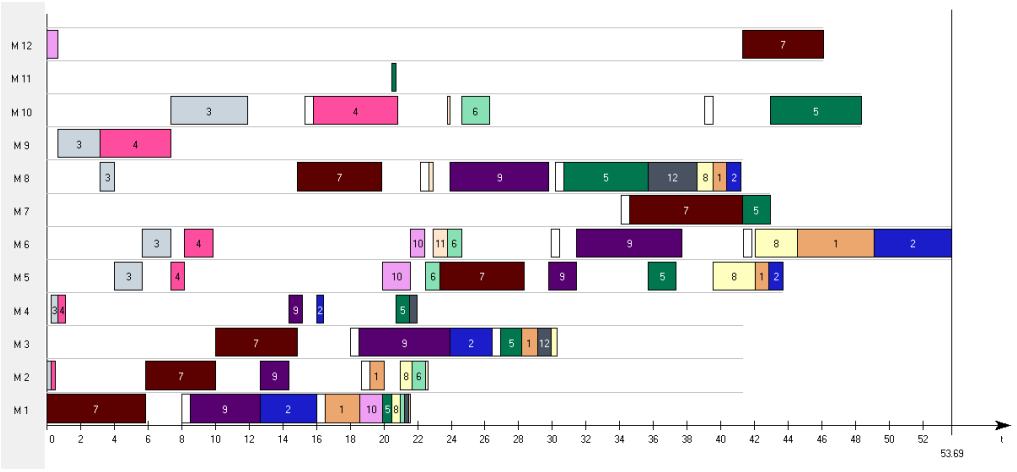
Przyjęto, iż wyroby produkowane są w partiach po 50 sztuk, a kryterium celu był termin zakończenia wszystkich zadań produkcyjnych – wskaźnik C_{max}.

Narzędziem harmonogramowania zadań było oprogramowanie LiSA będące zbiorem powszechnie stosowanych algorytmów szeregowania zadań i umożliwiające budowanie harmonogramów w typowych środowiskach produkcyjnych (*flow-shop*, *job-shop* czy *open-shop*) wraz z uwzględnieniem wybranych ograniczeń i kryteriów oceny [7]. Na rys. 10 przestawiono przykładowy harmonogram uzyskany w wyniku działania reguły LPT.



Rys. 10. Harmonogram nominalny – działanie reguły LPT

W celu uwzględnienia potencjalnych awarii maszyn technologicznych dokonano modyfikacji harmonogramów poprzez implementację buforów serwisowych o wielkości 0,5 godziny, mających charakter zabezpieczeń czasowych w przypadku wystąpienia awarii oraz czasu na dokonanie niezbędnej inspekcji, bądź działań serwisowych. Bufory implementowano zgodnie z czasami wyznaczonymi w wyniku realizacji algorytmu (tabela 2). Przyjęto, iż jest to wyłącznie czas pracy maszyny (realizacji zadań), po jakim może wystąpić usterka. Jeżeli w danym miejscu harmonogramu występowała operacja technologiczna – była ona przesuwana w prawo (zaraz za bufor) przy jednoczesnym zachowaniu kolejności zadań na jaki wskazywał harmonogram nominalny. Przykład harmonogramu odpornego z zaimplementowanymi buforami serwisowymi został przedstawiony na rysunku 11 (bufory oznaczono w postaci białych bloków).



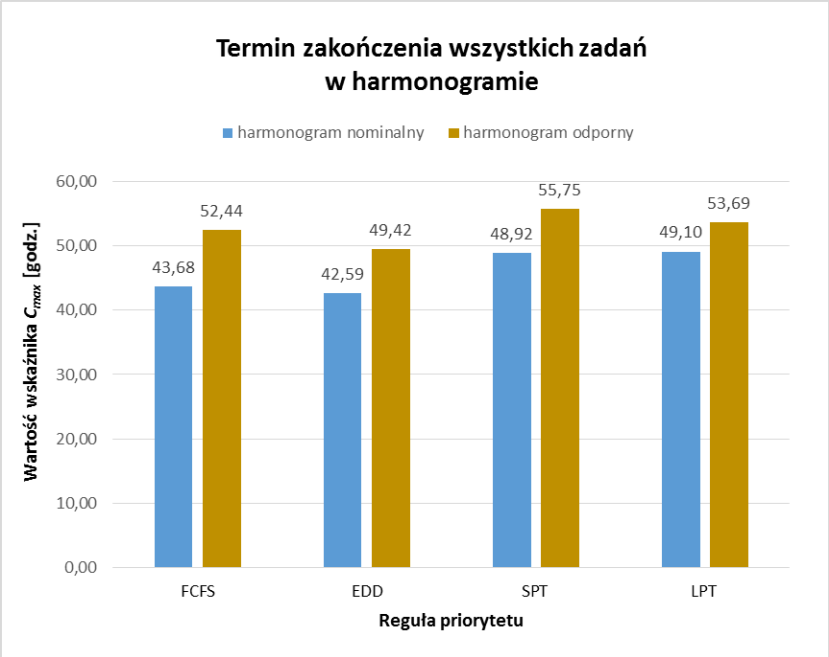
Rys. 11. Harmonogram produkcji po implementacji buforów serwisowych

Uzyskane czasy zakończenia wszystkich zadań produkcyjnych w harmonogramach nominalnych oraz odpornych przedstawiono w tabeli 3.

Tab. 3. Uzyskane wartości wskaźnika C_{max}

Reguła priorytetu	**Termin zakończenia zadań produkcyjnych – wskaźnik C_{max}** [godz.]		
	harmonogram nominalny	harmonogram odporny	wydłużenie [%]
FCFS	43,68	52,44	16,7%
EDD	42,59	49,42	13,8%
SPT	48,92	55,75	12,3%
LPT	49,10	53,69	8,5%

Uzyskane podczas eksperymentu wartości terminów zakończenia wszystkich zleceń po implementacji buforów serwisowych spowodowały wydłużenie kryterium celu dla każdego z przypadków. Średnia różnica pomiędzy wskaźnikiem C_{max} harmonogramu nominalnego i odpornego wyniosła 6,75 godz. Można zatem stwierdzić, iż uwzględnienie aspektu występowania awarii maszyn powoduje, iż produkcja będzie realizowana o około jedną zmianę dłużej. Procentowe wydłużenia analizowanego wskaźnika były odmienne dla harmonogramów uzyskanych z wykorzystaniem różnych reguł priorytetu. Przyjmowały one wartości od 8,5% dla harmonogramu opracowanego zgodnie z regułą LPT do 16,7% dla harmonogramu opracowanego z wykorzystaniem reguły FCFS. Uzyskane wartość terminów realizacji wszystkich zadań dla poszczególnych harmonogramów zostały zestawione na rysunku 12.

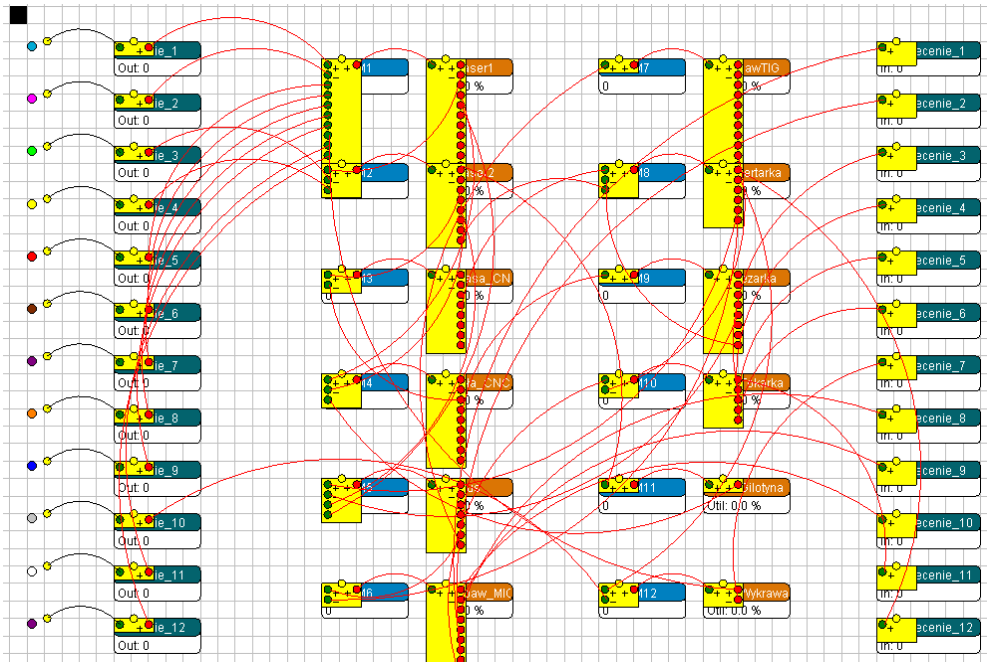


Rys. 12. Wartości wskaźnika C_{max} – terminu zakończenia wszystkich zadań

W celu oceny zasadności implementacji buforów, a w konsekwencji wydłużenia harmonogramu produkcji przeprowadzono drugą część eksperymentu polegającą na symulacji produkcji w warunkach niepewności. Pozwolił on wskazać, który z harmonogramów – nominalny czy odporny (uzyskany z wykorzystaniem proponowanego algorytmu) – wskazuje bliższy realnemu termin zakończenia wszystkich zadań produkcyjnych.

3.3.2. Symulacja produkcji w warunkach występowania awarii maszyn

Drugi etap eksperymentu został zrealizowany w środowisku symulacyjnym Enterprise Dynamics, którego obszary zastosowań obejmują: produkcję, magazynowanie, zarządzanie łańcuchem dostaw, systemy transportu i inne. Umożliwia ono zarówno modelowanie, symulację, jak i wizualizację procesów, co czyni z niego narzędzie pomocne w procesie kontroli procesów dynamicznych [14, 16, 22]. Za pomocą dostępnych elementów środowiska wykonano model pozwalający analizować realizację produkcji w rozpatrywanym systemie wytwórczym (rys. 13).



Rys. 13. Model systemu produkcyjnego opracowany w środowisku ED

W celu uwzględnienia awaryjności maszyn technologicznych dla każdej z nich (we właściwościach danego bloku) zdefiniowano wartości wskaźników MTTF oraz MTTR. Wartości parametru MTTF zostały zdefiniowane z wykorzystaniem jednostajnego rozkładu prawdopodobieństwa tak, aby awarie występowały w dowolnym czasie – od chwili rozpoczęcia pracy na danej maszynie, aż do chwili jej zakończenia. Parametr MTTR został określony z wykorzystaniem rozkładu gamma, gdyż właśnie taki rozkład został zidentyfikowany oraz dopasowany podczas analizy statystycznej posiadanych danych historycznych dotyczących czasów napraw maszyn. Przyjęte parametry wskaźników MTTF oraz MTTR dla poszczególnych maszyn zostały przedstawione w tabeli 4. Czas stanowiące parametry rozkładów zostały wyrażone w godzinach.

Tab. 4. Czasy awarii maszyn technologicznych uzyskane w wyniku predykcji

	Parametry wskaźników efektywności	
Maszyna	**MTTF**	**MTTR**
M_1	Uniform(0, 18,42)	Gamma(0,85, 1,62)
M_2	Uniform(0, 8,0)	Gamma(0,75, 2,07)
M_3	Uniform(0, 16,08)	Gamma(0,69, 2,79)
M_4	Uniform(0, 3,33)	Gamma(0,77, 1,88)
M_5	–	–
M_6	Uniform(0, 23,74)	Gamma(0,95, 2,43)
M_7	Uniform(0, 8,34)	Gamma(0,93, 1,96)
M_8	Uniform(0, 22,41)	Gamma(0,66, 2,45)
M_9	Uniform(0, 6,67)	Gamma(0,80, 1,64)
M_{10}	Uniform(0, 16,84)	Gamma(0,72, 1,85)
M_{11}	Uniform(0, 0,21)	Gamma(0,88, 2,16)
M_{12}	Uniform(0, 5,5)	Gamma(0,84, 1,78)

W opracowanym modelu modyfikowano także kolejność realizacji zadań na poszczególnych maszynach (tak, aby produkcja odbywała się zgodnie z harmonogramami opracowanymi z wykorzystaniem reguł: FCFS, EDD, SPT oraz LPT). W sumie badaniu poddano 4 modele, a w każdym z nich zrealizowano 25 symulacji produkcji.

Podczas oceny rezultatów przeprowadzonych badań symulacyjnych wykorzystano następujące wskaźniki stabilności:

– wskaźnik odchylenia terminu zakończenia wszystkich zadań $\Delta C_{\max}$ określany jako:

$$\Delta C_{\max} = C_{\max} - C'_{\max}, \tag{11}$$

gdzie: $\Delta C_{\max}$ – odchylenie terminu zakończenia wszystkich zadań,
$C_{\max}$ – długość harmonogramu bazowego,
$C'_{\max}$ – długość harmonogramu aktualnego (zrealizowanego).

– wskaźnik względnego wydłużenia terminu zakończenia wszystkich prac $E_{C\max}$, określanego na podstawie zależności:

$$E_{C_{\max}} = \frac{C_{\max}}{C'_{\max}}, \tag{12}$$

gdzie: $E_{C\max}$ – wskaźnik względnego wydłużenia terminu zakończenia wszystkich prac.

W tabeli 5 przedstawiono wyniki symulacji dla przypadku, gdy zadania produkcyjne uszeregowane były zgodnie z działaniem reguły SPT. Dla każdej z symulacji uzyskane wartości wskaźników stabilności potwierdziły skuteczność oraz zasadność wykorzystania proponowanego algorytmu. Zarówno wartości odchylenia terminu zakończenia wszystkich zadań $\Delta C_{\max}$, jak i względnego wydłużenia terminu zakończenia wszystkich $E_{C\max}$ wykazały, iż harmonogram uwzględniający potencjalne awarie maszyn technologicznych wskazuje bliższy realnemu termin zakończenia wszystkich zadań produkcyjnych.

Tab. 5. Wartości wskaźników stabilności – kolejność zadań zgodnie z regułą SPT

Nr sym.	**Harmonogram zrealizowany (symulacja)** C'_{max} [godz.]	**Odchylenie terminu zakończenia oraz względne wydłużenia terminu zakończenia**					
		harmonogram nominalny			harmonogram odporny		
		C_{max} [godz.]	ΔC_{max} [godz.]	E_{Cmax} [–]	C_{max} [godz.]	ΔC_{max} [godz.]	E_{Cmax} [–]
1	56,10		-7,18	0,87		-0,35	0,99
2	53,88		-4,96	0,91		1,87	1,03
3	54,09		-5,17	0,90		1,66	1,03
4	56,91		-7,99	0,86		-1,16	0,98
5	52,60		-3,68	0,93		3,15	1,06
6	55,50		-6,58	0,88		0,25	1,00
7	56,43		-7,51	0,87		-0,68	0,99
8	55,88		-6,96	0,88		-0,13	1,00
9	53,48		-4,56	0,91		2,27	1,04
10	54,04		-5,12	0,91		1,71	1,03
11	58,31		-9,39	0,84		-2,56	0,96
12	52,97	**48,92**	-4,05	0,92	**55,75**	2,78	1,05
13	54,20		-5,28	0,90		1,55	1,03
14	55,33		-6,41	0,88		0,42	1,01
15	55,98		-7,06	0,87		-0,23	1,00
16	56,01		-7,09	0,87		-0,26	1,00
17	53,53		-4,61	0,91		2,22	1,04
18	56,51		-7,59	0,87		-0,76	0,99
19	55,18		-6,26	0,89		0,57	1,01
20	56,49		-7,57	0,87		-0,74	0,99
21	52,37		-3,45	0,93		3,38	1,06
22	57,52		-8,60	0,85		-1,77	0,97
23	54,86		-5,94	0,89		0,89	1,02
24	55,04		-6,12	0,89		0,71	1,01
25	54,83		-5,91	0,89		0,92	1,02

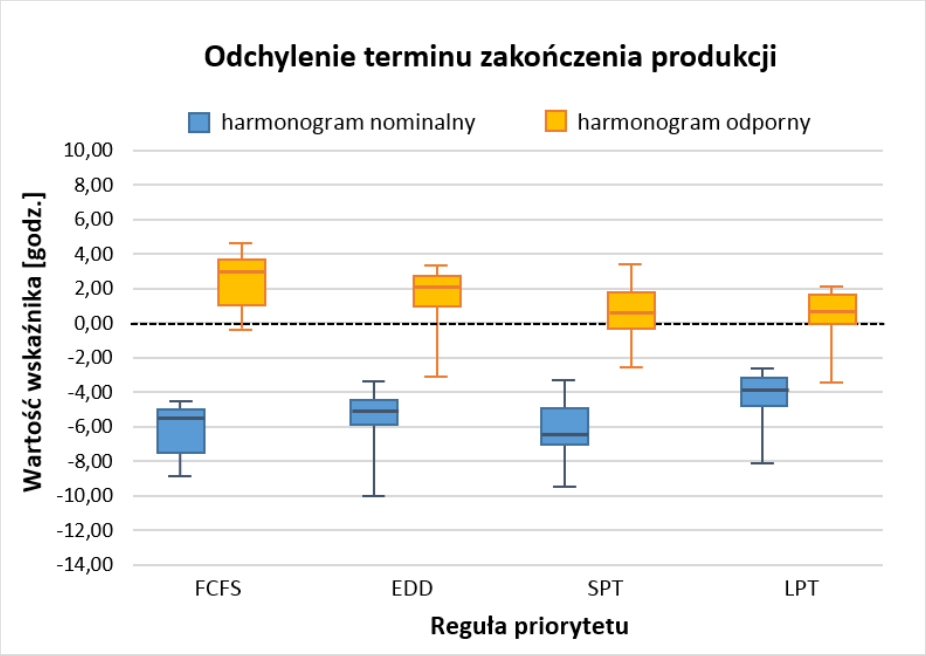
W przypadku pozostałych symulacji również została potwierdzona zasadność proponowanych w niniejszej publikacji rozwiązań. Świadczą o tym wartości średnie z poszczególnych symulacji zestawione w tabeli 6.

Tab. 6. Wartości średnie rozpatrywanych wskaźników stabilności

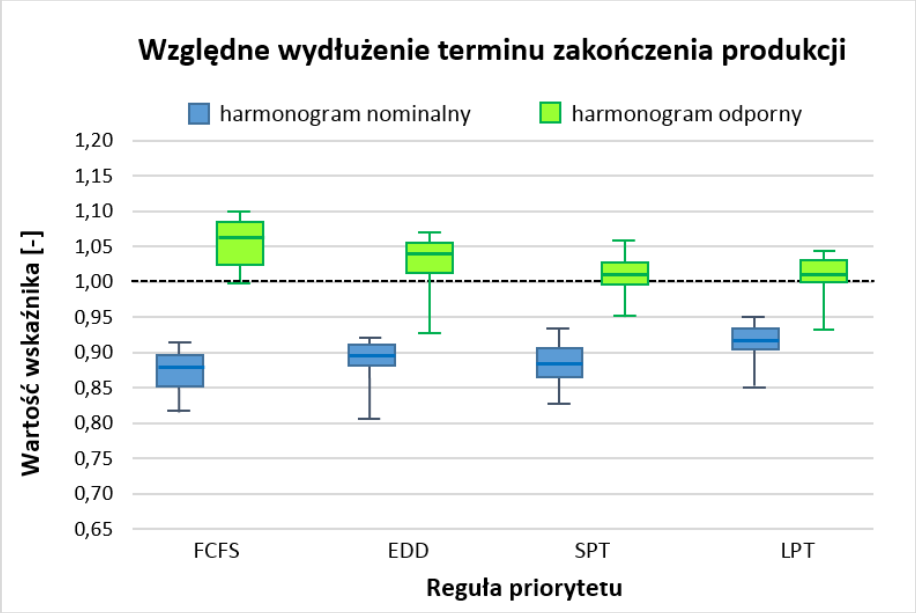
Reguła priorytetu	**Harmonogram zrealizowany (symulacja)** $\overline{C'}_{max}$ [godz.]	**Odchylenie terminu zakończenia oraz względne wydłużenia terminu zakończenia**					
		harmonogram nominalny			harmonogram odporny		
		C_{max} [godz.]	$\Delta\overline{C}_{max}$ [godz.]	$\overline{E}_{C_{max}}$ [–]	C_{max} [godz.]	$\Delta\overline{C}_{max}$ [godz.]	$\overline{E}_{C_{max}}$ [–]
FCFS	49,87	43,68	-6,19	0,88	52,44	2,57	1,05
EDD	47,90	42,59	-5,31	0,89	49,42	1,52	1,03
SPT	55,12	48,92	-6,20	0,89	55,75	0,63	1,01
LPT	53,14	49,10	-4,04	0,92	53,69	0,55	1,01

Uzyskane wartości wskazują wyraźnie, iż harmonogram z zaimplementowanymi buforami serwisowymi wskazywał bliższy realnemu termin zakończenia produkcji.

Na rysunku 14 oraz 15 zestawiono uzyskane wartości rozpatrywanych wskaźników, które również potwierdzają zasadność wykorzystania proponowanego algorytmu.



Rys. 14. Wartość wskaźnika odchylenia terminu zakończenia wszystkich zadań ΔC_{max}



Rys. 15. Wartości wskaźnika względnego wydłużenia terminu zakończenia wszystkich prac E_{Cmax}

Wykorzystanie opracowanego algorytmu pozwala wskazać bliższy realnemu termin zakończenia produkcji w warunkach, gdy istnieje ryzyko wystąpienia awarii maszyn technologicznych. Świadczy o tym chociażby fakt, iż dla harmonogramu odpornego wartości wskaźnika E_{Cmax} koncentrują się w okolicy wartości 1, zaś wartość wskaźnika ΔC_{max} w okolicy 0 – oznacza to dużą zgodność terminów zakończenia produkcji w harmonogramach odpornych z terminami uzyskanymi w wyniku symulacji produkcji.

4. Podsumowanie i wnioski końcowe

Predykcja awarii maszyn jest tematem wielu publikacji naukowych. Autorzy liczny prac starają się implementować różnorodne metody w celu wyznaczania informacji dotyczących awaryjności maszyn technologicznych. Wiarygodne i dobrze opracowane plany prac prewencyjnych stanowią kluczowy element działań związanych z utrzymaniem ruchu, szczególnie w obszarze wykorzystania strategii *Time-Based Maintenance*.

W niniejszej pracy zaprezentowano algorytm predykcji zorientowany na wykorzystanie typowych danych historycznych, posiadane przez działy UR. Proponowany algorytm stanowi alternatywne podejście do problemu predykcji awarii, bowiem wykorzystanie estymacji Kaplana-Meier'a pozwala na wyznaczenie charakterystyk występowania awarii w czasie dla poszczególnych maszyn technologicznych systemu wytwórczego, co w konsekwencji wspomaga działania TBM. Zastosowanie elementów analizy czasu trwania powoduje, iż przedstawione rozwiązanie jest innowacyjnym oraz konkurencyjny w zakresie wnioskowania na podstawie rzeczywistych danych historycznych. W konsekwencji kluczowym aspektem staje się zatem gromadzenie rzetelnych danych dotyczących awaryjności maszyn. Tylko odpowiednia ilość oraz jakoś danych historycznych pozwala uzyskać wiarygodne i miarodajne rezultaty.

Opracowany algorytm wpisuje się w tendencję coraz szerszego wykorzystania narzędzi informatycznych w pracach działów UR. Dlatego też został on opracowany w taki sposób, aby możliwa była jego implementacja w postaci programu komputerowego, bądź dodatku do znanych już rozwiązań. Weryfikacja proponowanego algorytmu pozwoliła wyznaczyć potencjalne czasy występowania awarii maszyn technologicznych. Należy zauważyć, iż dla analizowanych maszyn czasy te były różne, co oznacza, iż każda z nich posiada własną charakterystykę występowania usterek. Potwierdza to słuszność oraz potrzebę wykorzystania strategii TBM w procesie utrzymywania ruchu obiektów technicznych. Uzyskane informacje są również niezwykle istotne w aspekcie realizacji produkcji w warunkach niepewności. Przeprowadzone w drugiej części publikacji badania symulacyjne dowodzą, iż wykorzystanie rezultatów proponowanego algorytmu w procesie planowania produkcji, pozwala uzyskać stabilność realizowanych procesów, a w konsekwencji wskazywać bliższe realnemu terminy zakończenia produkcji.

Zrealizowane badania potwierdzają skuteczność opracowanego algorytmu predykcji, a także wskazują na potrzebę realizacji działań prewencyjnych w kierunku zapobiegania występowaniu awarii maszyn w celu zapewnienia większej stabilności realizowanych procesów.

Projekt/Badania zostały sfinansowane z Projektu Politechnika Lubelska – Regionalna Inicjatywa Doskonałości ze środków Ministerstwa Nauki i Szkolnictwa Wyższego na podstawie umowy nr 030/RID/2018/19.

References

1. Albrice D, Branch M. A Deterioration Model for Establishing an Optimal Mix of Time-Based Maintenance (TbM) and Condition-Based Maintenance (CbM) for the Enclosure System. Fourth Building Enclosure Science & Technology Conference (BEST4), Kansas City, Missouri, April 13–15, 2015.
2. Al-Hinai N, ElMekkawy TY. Robust and Stable Flexible Job Shop Scheduling with Random Machine Breakdowns Using a Hybrid Genetic Algorithm. International Journal of Production Economics 2011, 132(2): 279–291, http://dx.doi.org/10.1016/j.ijpe.2011.04.020.

3. Antosz K, Stadnicka D. Evaluation measures of machine operation effectiveness in large enterprises: study results. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2015; 17(1): 107–117, http://dx.doi.org/10.17531/ein.2015.1.15.
4. Baptista M, Sankararaman S, de Medeiros IP, Nascimento C, Prendinger H, Henriques EMP. Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling, Computers & Industrial Engineering 2018, 115: 41–53, https://doi.org/10.1016/j.cie.2017.10.033.
5. Bartochowska D, Ferenc R. Instrumenty wsparcia utrzymania ruchu w małych i średnich przedsiębiorstwach. Zeszyty naukowe Politechniki Śląskiej 2015, 80: 21–50.
6. Bei XQ, Zhu XY, Coit DW. A risk-averse stochastic program for integrated system design and preventive maintenance planning. European Journal Of Operational Research 2019, 276(2): 536–548, http://dx.doi.org/10.1016/j.ejor.2019.01.038.
7. Bräsel H, Dornheim L, Kutz S, Mörig M, Rössling I. LiSA – A Library of Scheduling Algorithms. Magdeburg University, 2001.
8. Davenport A, Gefflot C, Beck C. Slack-based Techniques for Robust Schedules. Sixth European Conference on Planning, Toledo, Spain, September 12–14, 2001.
9. Deepu P. Robust Schedules and Disruption Management for Job Shops. Bozeman, Montana, 2008.
10. Fernandes M, Canito A, Bolon-Canedo V, Conceicao L, Praca I, Marreiros G. Data analysis and feature selection for predictive maintenance: A case-study in the metallurgic industry. International Journal Of Information Management 2019, 45: 252–262, http://dx.doi.org/10.1016/j.ijinfomgt.2018.10.006.
11. Frątczak E, Sienkiewicz U, Babiker H. Analiza historii zdarzeń – Elementy teorii, wybrane przykłady zastosowań. Oficyna Wydawnicza Szkoła Główna Handlowa w Warszawie, Warszawa 2014.
12. Gao H. Bulding Robust Schedules using Temporal Potection – An Empirical Study of Constraint Based Scheduling Under Machine Failure Uncertainty. Toronto, Ontario, 1996.
13. Gao Y, Feng Y, Zhang Z, Tan J. An optimal dynamic interval preventive maintenance scheduling for series systems. Reliability Engineering & System Safety 2015, 142: 19–30, http://dx.doi.org/10.1016/j.ress.2015.03.032.
14. Gola A. Reliability analysis of reconfigurable manufacturing structures using computer simulation methods. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21(1): 90–102, http://dx.doi.org/10.17531/ein.2019.1.11.
15. Gürel S, Körpeoḡlu E, Aktürk MS. An Anticipative Scheduling Approach with Controllable Processing Times. Computers & Operations Research 2010, 37(6): 1002–1013, http://dx.doi.org/10.1016/j.cor.2009.09.001.
16. Jasiulewicz-Kaczmarek M, Bartkowiak T. Improving the performance of a filling line based on simulation, ModTech International Conference – Modern Technologies in Industrial Engineering IV, Romania, Iasi, June 15–18, IOP Conf. Series: Materials Science and Engineering 2016, 145(042024), https://doi.org/10.1088/1757-899X/145/4/042024.
17. Jensen MT. Improving robustness and flexibility of tardiness and total flow-time job shops using robustness measures. Applied Soft Computing 2001, 1: 35–52, http://dx.doi.org/10.1016/S1568-4946(01)00005-9.
18. Jian X, Li-Ning X, Ying-Wu Ch. Robust Scheduling for Multi-Objective Flexible Job-Shop Problems with Random Machine Breakdowns. International Journal of Production Economics 2013, 141(1): 112–126. https://doi.org/10.1016/j.ijpe.2012.04.015.

19. Kalinowski K, Krenczyk D, Grabowik C. Predictive-reactive strategy for real time scheduling of manufacturing systems. Applied Mechanics and Materials 2013, 307: 470–473, https://doi.org/10.4028/www.scientific.net/AMM.307.470.
20. Kempa W, Paprocka I, Kalinowski K, Grabowik C. Estimation of reliability characteristics in a production scheduling model with failures and time-changing parameters described by Gamma and exponential distributions. Advanced Materials Research 2014, 837: 116–121.
21. Kempa W, Wosik I, Skołud B. Estimation of Reliability Characteristics in a Production Scheduling Model with Time-Changing Parameters – First Part, Theory. Management and Control of Manufacturing Processes. Lublin, 2011: 7–18.
22. Kłos S, Patalas-Maliszewska J, Trebuna P. Improving manufacturing processes using simulation methods. Applied Computer Science 2016, 12(4): 7–17.
23. Lawless J. F. Statistical Models and Methods for Lifetime Data. John Wiley & Sons, 2003.
24. Leon VJ., Wu SD., Storer RH. Robustness Measures and Robust Scheduling for Job Shops. IIE transactions 1994, 26(5): 32–43, https://doi.org/10.1080/07408179408966626.
25. Liao W, Zhang X, Jiang M. An optimization model integrated production scheduling and preventive maintenance for group production. IEEE International Conference on Industrial Engineering and Engineering Management 2016, December, 936–940, http://dx.doi.org/10.1109/IEEM.2016.7798015.
26. Loska A. Scenario modeling exploitation decision-making process in technical network systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (2): 268–278, http://dx.doi.org/10.17531/ein.2017.2.15.
27. Lü Y, Zhang Y. Reliability Modeling and Maintenance Policy Optimization for Deteriorating System Under Random Shock. Journal of Shanghai Jiaotong University (Science) 2018, 23(6): 791–797, http://dx.doi.org/10.1007/s12204-018-1985-y.
28. Mehta SV., Uzsoy RM. Predictable Scheduling of a Job Shop Subject to Breakdowns. IEEE Transactions on Robotics and Automation 1998, 14(3): 365–378, https://doi.org/10.1109/70.678447.
29. Rawat M, Lad BK., Novel approach for machine tool maintenance modelling and optimization using fleet system architecture. Computers & Industrial Engineering 2018, 126: 47–62, http://dx.doi.org/10.1016/j.cie.2018.09.006.
30. Rosmaini A, Shahrul K. An overview of time-based and condition-based maintenance in industrial application. Computers & Industrial Engineering 2012; 63(1): 135–149, http://dx.doi.org/10.1016/j.cie.2012.02.002.
31. Sabuncuoglu I, Bayõz M. Analysis of reactive scheduling problems in a job shop environment. European Journal of Operational Research 2000, 126(3): 567–586, https://doi.org/10.1016/S0377-2217(99)00311-2.
32. Skołud B., Wosik I., Immune Algorithms in Production Jobs Scheduling. Zarządzanie Przedsiębiorstwem 2008, 1: 47–48.
33. Sobaszek Ł, Gola A, Kozłowski E. Job-shop scheduling with machine breakdown prediction under completion time constraint. Annals of Computer Science and Information Systems 2018; 15: 437–440, http://dx.doi.org/10.15439/2018F83.
34. Szwedzka K, Szafer P, Wyczółkowski R. Structural analysis of factors affecting the effectiveness of complex technical systems. Proceedings of the 30th International Business Information Management Association Conference, IBIMA 2017 – Vision 2020: Sustainable Economic development, Innovation Management, and Global Growth Volume 2017, 4096–4105.

35. Timofiejczuk A, Brodny J, Loska A. Exploitation Policy in the Aspect of Industry 4.0 Concept – Overview of Selected Research. Multidisciplinary Aspects of Production Engineering 2018, 1(1): 353–359. https://doi.org/10.2478/mape-2018-0045.
36. Vonta F. Frailty or Transformation Models in Survival Analysis and Reliability. Recent Advances In System Reliability: Signatures, Multi-State Systems And Statistical Inference 2012, 237–251, http://dx.doi.org/10.1007/978-1-4471-2207-4_17.
37. Wei-Wei C, Zhiqiang L, Ershun P. Integrated Production Scheduling and Maintenance Policy for Robustness in a Single Machine. Computers & Operations Research 2014, 47: 81–91, https://doi.org/10.1016/j.cor.2014.02.006.
38. Yang BY, Liu RN, Zio E. Remaining Useful Life Prediction Based on a Double-Convolutional Neural Network Architecture. IEEE Transactions On Industrial Electronics 2019, 66(12): 9521–9530, https://doi.org/10.1109/TIE.2019.2924605.
39. Zhang F, Shen J, Ma Y. Optimal maintenance policy considering imperfect repairs and non-constant probabilities of inspection errors. Reliability Engineering and System Safety 2020, 193, http://dx.doi.org/10.1016/j.ress.2019.106615.
40. Zhao X, He S, He Z, Xie M. Optimal condition-based maintenance policy with delay for systems subject to competing failures under continuous monitoring. Computers & Industrial Engineering 2018, 124: 535–544, http://dx.doi.org/10.1016/j.cie.2018.08.006.

dr inż. Robert Konowrocki
Instytut Podstawowych Problemów Techniki, PAN
ul. Pawińskiego 5B, 02-106 Warszawa, Polska

Instytut Kolejnictwa
ul. Chłopickiego 50, 04-275 Warszawa, Polska
e-mail: rkonow@ippt.pan.pl

mgr inż. Andrzej Chojnacki
Instytut Kolejnictwa
ul. Chłopickiego 50, 04-275 Warszawa, Polska
e-mail: achojnacki@ikolej.pl

Analiza niezawodności eksploatacyjnej pojazdów szynowych w aspekcie bezpieczeństwa przed wykolejeniem w oparciu o różne metody wyznaczania kryterium oceny

Słowa kluczowe: bezpieczeństwo eksploatacji, dynamika pojazdów szynowych, wykolejenie, badania numeryczne i eksperymentalne

Streszczenie: W pracy pokazano rezultaty badań komputerowych i eksperymentalnych dotyczących zagadnień eksploatacji w aspekcie bezpieczeństwa w odniesieniu do wykolejenia wagonu towarowego na torze kolejowym. Przybliżono w nim stan wiedzy dotyczącej metod oceny bezpieczeństwa eksploatacji pojazdówa) szynowych na kolejowych liniach szynowych, w celu ich analizy porównawczej. W pracy wykonano analizy teoretyczne bazując na kilku metodach, które oceniają bezpieczeństwo ich wykolejenia, kwalifikujące się do niezawodność eksploatacyjnej, porównując je z wynikami otrzymanymi z badań eksperymentalnych. Na potrzeby przeprowadzanych badań powstał komputerowy model pojazd szynowy - tor kolejowy. Uwzględniał on parametry dynamiczne elementów zastosowanych w rzeczywistym torze oraz pojeździe szynowym. Otrzymane z teoretycznych analiz wyniki zwalidowano testami eksperymentalnymi wykonanymi na rzeczywistych obiektach (pojazd towarowy - tor testowy, wagon towarowy - stanowisko badawcze). W ramach badań zaproponowano nową geometrię toru testowego do badań pojazdów szynowych. Uzyskane wyniki pozwoliły określić stan zagrożenia eksploatacji wagonu towarowego podczas jazdy po testowej infrastrukturze szynowej przy różnych kryteriach oceny oraz je porównać.

1. Wstęp

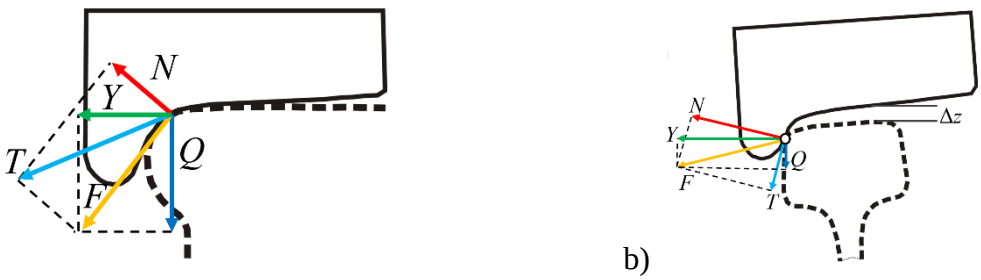
Problematyka bezpieczeństwa i niezawodności ruchu pojazdów szynowych w badaniach naukowych stale jest rozwijana [41]. Procesy zmierzające do podniesienia poziomu niezawodności eksploatacyjnej oraz bezpieczeństwa uwzględniane są już na etapie ich projektowania. Światowe trendy dotyczące tych kwestii wymagają wykorzystywania teorii niezawodności bezpieczeństwa i niezawodności funkcjonowania takich pojazdów [29, 34, 39]. Teoria bezpieczeństwa powstała w latach dziewięćdziesiątych, aby przeciwdziałać ryzykom awarii i wypadków, które mogą prowadzić nie tylko do przerwania funkcjonowania danego systemu technicznego, ale także utraty zdrowia i życia ludzi lub innych szkód [21]. W przypadku pojazdów szynowych wykolejenie jest najczęstszym rodzajem wypadków kolejowych wywołujących jednocześnie ryzyko degradacji infrastruktury, uszkodzenia taboru, przewożonego towaru, zakłócenia świadczenia usług, a także szkody w środowisku (przewóz

niebezpiecznych materiałów) [16, 18]. Wobec powyższego wyznaczenie ryzyka wystąpienia pod względem tego zjawiska są kluczowe podczas badań pojazdów szynowych. Na etapie projektowania takich pojazdów prowadzone są analizy i badania teoretyczne prognozujące wpływ parametrów pojazdu na oddziaływanie z torem [1, 9, 28] oraz bezpieczeństwo, monitoring i niezawodność podczas jazdy związanej z wykolejeniem [6, 7]. Kontynuuje się je doświadczalnie jako badania dopuszczeniowe podczas kwalifikacji pojazdu do eksploatacji oraz po znaczących naprawach/modernizacjach eksploatowanych pojazdów szynowych [11, 35]. Bezpieczeństwo przed wykolejeniem jest jednym z zasadniczych kryterium oceny niezawodności ruchu pojazdów szynowych [40].

Wielu badaczy wciąż zajmuje się tematyką bezpieczeństwa pojazdów szynowych [10, 48, 49]. W wielu przypadkach głównym mechanizmem wywołującym wykolejenia pociągu jest utrata stabilności bocznej pojazd kolejowego [6, 24, 43]. Wywołuje to wzrostem wartości siły poprzecznej w strefie kontaktu koło-szyna. Może to być wynikiem różnych warunków skutkujących utratą bocznego prowadzenia, zapewnionego przez tor podczas normalnej pracy pojazdu. Należy tu wymienić: wznoszenie się obrzeża koła, poszerzenie rozstawu szyn, pochylenie szyn, stanu toru [49, 50] oraz zmniejszenie sztywność poprzecznej systemu przytwierdzenia do podkładów [20]. Oprócz badań teoretycznych prowadzone są także badania eksperymentalne wykolejenia wagonów towarowych [42, 50], w których analizowano dopuszczalne względne odciążenie koła zestawu kołowego $\lim\Delta q$. Wykazały one, że może się on zawierać w przedziale $0{,}62 \leq \lim\Delta q \leq 0{,}84$. W większości przypadków do analiz bezpieczeństwa służą kryteria oceny opierające się o wartości wskaźników wykolejenia Y/Q, odciążenia kół oraz ich wzniosu Δz. Na podstawie przeglądu literatury można dokonać podziału metodologii bezpieczeństwa przed wykolejeniem pojazdów szynowych przy różnych kryteriach oceny. Wymienić tutaj należy:

1. Kryterium graniczne Y/Q Nadala dla pojedynczego koła – obowiązujące dla małych prędkości przejazdu przez łuki toru [10].
2. Kryterium graniczne sumy osi Y/Q Weinstock [48].
3. Limit czasu CHXI 50 milisekund - Stowarzyszenie Kolei Amerykańskich, U.S. [12].
4. Kryterium czasu trwania Y/Q - zaproponowane przez Japońskie Koleje Państwowe [32].
5. Kryterium czasu trwania Y/Q - proponowane przez Dział Elektromotoryczny General Motors (EMD) [25].
5. Kryterium wysokości wznoszenia kół - zaproponowane przez Transportation Technology Center, Inc. (TTCI) [47].

Istnieje kilka przyczyn wystąpienia ryzyka wykolejenia pojazdu szynowego. Jeden z głównych scenariuszy wykolejenia jest realizowany, gdy podczas ruchu pojazdu duża siła boczna działająca na zestaw kołowy prowadzi do kontaktu obrzeża koła z szyną. W wyniku tego kontaktu koło szybko wspina się po szynie i po osiągnięciu maksymalnej wartości kąta obrzeża zestaw kołowy wykoleja się. Wznoszenie koła wywołujące wykolejenie jest związane z przekroczeniem wartości granicznej stosunku wartości składowych siły bocznej Y do siły pionowej Q na styku koło-szyna, patrz Rys. 1. W takim przypadku stosunek sił Y/Q jest zwykle nazywany współczynnikiem wykolejenia.



a) b)

Rys. 1. Składowe sił w kontakcie koło-szyna na prostej (a) i na łuku toru (b): siły poprzeczne (*Y*) i pionowa (*Q*), siła normalna (*N*), boczna siła tarcia tocznego (*F*), pochylenia obrzeża (*γ*), (*Δz*) wznios koła

Wartości składowych sił prezentowanych na rysunku 1 można określić z zależności (1)

$$\begin{aligned} Y &= Fcos(\gamma) + Nsin(\gamma) = Y_{cos} + Y_{sin} \\ Q &= -Fsin(\gamma) + Qscos(\gamma) = Q_{cos} + Q_{sin} \end{aligned} \quad . \tag{1}$$

Kryterium opracowane przez Nadala [10] określające współczynnik wykolejenia bazuje na wartości współczynnika tarcia μ pomiędzy kołem a szyną oraz kącie pochylenia obrzeża γ. Matematyczne określa je zależność (2). Kryterium to jest łatwe do wdrożenia i dlatego jest szeroko stosowane do oceny bezpieczeństwa przed wykolejeniem.

$$\frac{|Y|}{|Q|} < \frac{tg\gamma - \mu}{1 + \mu tg\gamma} \quad . \tag{2}$$

W szczególności powyższe kryterium jest stosowane do oceny ryzyka wykolejenia poprzez wzniesienie obrzeża w karcie UIC 518 [46] i normie europejskiej EN14363 [11]. Główną modyfikacją przyjętą w tych dwóch dokumentach jest wymóg, aby stosunek *Y*/*Q* nie przekroczył zakładaną wartość krytyczną równą 1,2 w odstępie 2 m dystansu przejazdu pojazdu w przypadku badań quasistatycznych.

Znany jest także inny znormalizowany [17] indeks stosowany do oceny dynamiki taboru związanego z bezpieczeństwem ruchu na torach o szerokości 1520 mm. Nazwany jest on współczynnikiem bezpieczeństwa stabilności zestawu kołowego przed wykolejeniem w przypadku toczenia kołnierza koła przy/po główce szyny i oznaczany przez k_z. Wyznacza się go z zależności (3). Współczynnik ten jest modyfikacją matematyczną zależności (2) i jego maksymalna dopuszczalna wartość dla wagonów towarowych wynosi 1,3 [17].

$$k_z < \frac{tg\beta - f_{FR}\mu}{1 + f_{FR} \cdot tg\beta} \cdot \frac{P_v}{Y} \geq [k_z] \ , \tag{3}$$

gdzie: β jest kątem nachylenia stożkowej części powierzchni kołnierza koła do poziomej linii odniesienia, f_{FR} opisuje współczynnik tarcia ślizgowego w trefie kontaktu koła i szyny, P_v i Y są odpowiednio pionową i poziomą składową sił oddziaływania koła na szynę.

Celem niniejszej pracy jest analiza oceny bezpieczeństwa a tym samym niezawodności eksploatacyjnej pojazdów szynowych pod względem wykolejenia, jak również propozycja nowej geometrii toru testowego, który uwzględnia wichrowatość toru na bazie wózka i bazie pojazdu. Taki tor umożliwia podczas jednego przejazdu pojazdu szynowego określenie wskaźników bezpieczeństwa przed wykolejeniem, bez potrzeby wykonywania dodatkowych testów na stacjonarnych stanowiskach opisanych w drugiej sekcji pracy.

2. Metody oceny ryzyka i bezpieczeństwa przed wykolejeniem stosowane podczas badań eksperymentalnych

Nowe konstrukcje pojazdów szynowych, które zostaną dopuszczone do eksploatacji na terytorium Unii Europejskiej, powinny spełniać zasadnicze wymagania podane w Technicznych Specyfikacjach Interoperacyjności (TSI). Jednym z wymagań jest sprawdzenie, czy dany pojazd może być bezpiecznie eksploatowany na torach. Zarówno TSI dotyczące wagonów towarowych, jak i TSI dotyczące lokomotyw i pojazdów pasażerskich, za dowód bezpieczeństwa uznaje spełnienie wymagań podanych w normie EN 14363:2005 (obecnie obowiązujące wydanie normy to EN 14363:2016 [11] – w Polsce przyjęta do stosowania i oznaczona jako PN-EN 14363+A1:2019-02). Przed ustanowieniem ww. normy badania ryzyka i bezpieczeństwa przed wykolejeniem wagonów towarowych wykonywano w oparciu o wymagania podane w raporcie ERRI (ORE) [40]. Dla oceny bezpieczeństwa jazdy po zwichrowanym torze innych pojazdów szynowych nie opracowano odrębnych wymagań. W związku z tym, raport [40] był wykorzystywany również do wyznaczenia współczynnika bezpieczeństwa jazdy po zwichrowanym torze tych pojazdów. Norma [11] odnosi się zarówno do badań quasistatycznych (prędkość badanego pojazdu szynowego nie przekracza 10 km/h), jak i badań dynamicznych (przewidziane dla pojazdów o dopuszczalnej prędkości powyżej 60 km/h). W omawianym dokumencie wymienione są 3 metody badania bezpieczeństwa jazdy po zwichrowanym torze.

Metoda 1 – badanie pojazdu podczas przejazdu przez zwichrowany tor. W tym przypadku jako stanowisko pomiarowe używany jest tor o stałym promieniu łuku R = 150 m. Wichrowatość toru realizowana jest poprzez zmianę wysokości położenia szyny zewnętrznej (przechyłka toru dodatnia i ujemna). Niezbędna w badaniu wichrowatość toru wynosi 3 ‰ na odcinku o długości 30 m. Dodatkowo budowa toru powinna odzwierciedlać normalne warunki typowego toru z uwzględnieniem profilu szyny, szerokości toru i stanu utrzymania. W trakcie badań nie mogą występować siły wzdłużne w składzie pociągu a sam pojazd nie może być hamowany. Przejazd przez łuk realizowany jest poprzez pchanie lub ciągnięcie pojazdu. Badania są prowadzone na szynach suchych, aby współczynnik tarcia między powierzchną koła i szynie był największy. Poprzedni dokument Raport ORE [40] dotyczący badań bezpieczeństwa i dynamiki jazdy wagonów towarowych po zwichrowanym torze zalecał, aby przed rozpoczęciem badań tor przemyć rozpuszczalnikiem technicznym. W kolejnym etapie przygotowań toru wymagane było posypanie szyny drobnym piaskiem, a następnie zmiecenie go z powierzchni główki szyny. Tak przygotowany tor zapewniał duży współczynnik tarcia między kołem a szyną.

Europejska norma [11] podaje zależności matematyczne, z których należy wyznaczyć wymaganą w badaniach bezpieczeństwa wichrowatość na bazie wózka i bazie pojazdu. W przypadku, gdy jest ona większa od 3 ‰ należy, w odpowiedni sposób, przygotować pojazd. Można to zrealizować np. za pomocą podkładek umieszczonych pod usprężynowaniem pojazdu. Wskazówki dotyczące obliczeń grubości podkładek i ich rozmieszczenia zamieszczono w załączniku normy EN 14363. Przed rozpoczęciem badań bezpieczeństwa należy wyznaczyć na pojeździe pionowe naciski kół. Następnie badanie należy tak zaplanować, aby koło o najmniejszym pionowym nacisku było kołem prowadzącym, czyli kołem atakującym szynę zewnętrzną na łuku pomiarowym. W przypadku nie spełnienia tej wytycznej, należy doprowadzić do takiego stanu umieszczając dodatkowe podkładek pod usprężynowaniem w innej lokalizacji.

W czasie badań wymagane jest wykonanie minimum 3 prób przejazdu pojazdu przez łuk ze stałą prędkością nie przekraczającą 10 km/h. Parametrami mierzonymi podczas nich są: siły prowadzące na wewnętrznym i zewnętrznym kole badanego pojazdu Y_i, Y_a, siły pionowego

nacisku na wewnętrznym i zewnętrznym kole badanego pojazdu Q_i, Q_a, kąt nabiegania zestawu prowadzącego α, uniesienie koła prowadzącego Δz w całym łuku. Wymienione parametry mogą być mierzone przez urządzenia umieszczone w torze lub na pojeździe. W przypadku, gdy urządzenia pomiarowe umieszczone są w torze to ich lokalizacja powinna znajdować się na zwichrowanym odcinku toru. Szczegółowe wytyczne można znaleźć w normie EN 14363:2016 [11], a przykład lokalizacji punktów pomiarowych na łuku toru testowego zamieszczono na rysunku 6. Uniesienie koła prowadzącego Δz należy rejestrować w sposób ciągły. Pomiary sił (Y_i, Q_i) na szynie wewnętrznej oraz kąt nabiegania zestawu prowadzącego α służą jedynie do weryfikacji współczynnika tarcia między powierzchnią koła a szyną. Na podstawie zmierzonych wartości sił pionowych Q_a i poprzecznych Y_a z poszczególnych punktów pomiarowych na łuku toru wyznaczono współczynniki bezpieczeństwa jazdy $(Y/Q)_a$ w poszczególnych przekrojach pomiarowych. Oceniano maksymalne wartość współczynnika bezpieczeństwa jazdy po zwichrowanym torze $(Y/Q)_{a,max}$. Pojazd uznaje się za bezpieczny, jeżeli spełniony jest warunek (4).

$$(Y/Q)_{a,\,max} \leq (Y/Q)_{lim}\,. \tag{4}$$

Zgodnie z kryterium Nadal'a, dla koła o kącie pochylenia obrzeża 70° i współczynniku tarcia koła o szynę $\mu = 0{,}36$ warunek ten wynosi czyli $(Y/Q)_{a,max} > (Y/Q)_{lim}$. W przypadku, gdy wartość graniczna $(Y/Q)_{lim}$ jest przekroczona sprawdzeniu należy poddać uniesienie koła nad główkę szyny. Jeżeli warunek $\Delta z_{max} \leq \Delta z_{lim}$ (gdzie: $\Delta z_{lim} = 0{,}005$ m) jest spełniony oznacza to, że pojazd faktycznie nie uległ wykolejeniu. Dlatego też pojazd może zostać uznany za bezpieczny jeżeli spełnia dodatkowo następujące warunki: kąt obrzeża nie przekracza 70° w żadnym położeniu profilu, należy udokumentować, że zewnętrzna szyna jest sucha i nie ma na niej pozostałości smaru i ciał obcych, badanie było przeprowadzone przynajmniej 3 razy i w każdym przypadku warunek $\Delta z_{max} \leq \Delta z_{lim}$ został spełniony.

Zatem, jako ostateczne kryterium oceny bezpieczeństwa przejazdu pojazdu przez tor zwichrowany jest spełnienie kryterium uniesienie koła prowadzącego $\Delta z_{max} \leq \Delta z_{lim}$.

Metoda 2 analizy ryzyka wykolejenia pojazdu szynowego dotyczy badań przeprowadzonych na stanowisku symulującym oddziaływanie na pojazd toru zwichrowanego oraz przejazd badanego pojazdu przez tor badawczy bez wichrowatości. Do oceny bezpieczeństwa jazdy po zwichrowanym torze, w oparciu o metodę 2 - opisaną w normie [11], należy użyć dwóch stanowisk badawczych. Pierwsze to specjalne stanowiska do wichrowania, na którym to zostaje wyznaczony minimalny nacisk koła Q_a podczas przejazdu przez tor zwichrowany. Drugie stanowisko to tor o promieniu R = 150 m bez przechyłki. Na tym stanowisku jest wyznaczana maksymalna siła prowadząca Y_a zestaw kołowy. Za podstawę do wyznaczenia granicznych wichrowatości używane są te same zależności, które opisano w metodzie 1. Z uwagi na fakt, że w rzeczywistości dopuszczalne jest uniesienie koła do Δz = 0,005 m, zależności na graniczne wichrowatości, w badaniu stanowiskowym, są zredukowane.

Stanowisko używane do pomiarów nacisków kół Q_{jk} (j- oznacza numer zestawu kołowego, k- stronę pojazdu) musi być wyposażone w urządzenia służące do ich unoszenia i opuszczania. Niezależne przemieszczenie kół powinno być realizowane przynajmniej na dwóch zestawach kołowych jednego wózka. W trakcie wichrowania mierzone jest przemieszczenie Δz_{jk} kół w sposób ciągły oraz siły nacisku Q_{jk} wszystkich kół. Na podstawie obróbki danych, uwzględniając siły spowodowane łącznym wichrowaniem nadwozia i wózka, niecentrycznością ich środków ciężkości, włączając w to tarcie i odchyłki, zostaje wyznaczona minimalna siła pionowego nacisku koła $Q_{jk,min}$. Na Rys 2. przedstawiono przykładowe

przemieszczenie kół podczas wichrowania pojazdu, wykonywane w celu określenia wartości sił nacisku poszczególnych kół na toki szynowe. Proces ten wykonywany na stanowisku symuluje zmianę nacisków kół podczas przejazdu pojazdu przez tor zwichrowany.

a) 

b) 

Rys. 2. Pozycje wahaczy modułów pomiarowych podczas przeprowadzanych badań siły nacisku kół wagonu towarowego – wichrowanie wagonu towarowego, (a) przemieszczenie koła w górę względem zerowego poziomu główki szyny (b) przemieszczenie koła w dół względem zerowego poziomu główki szyny

W przypadku wyznaczania maksymalnej siły prowadzącej Y_a tor powinien składać się z odcinka prostego i krzywizny o promieniu R = 150 m. Stanowisko pomiarowe nie powinno posiadać krzywej przejściowej, przechyłki i wichrowatości. Podobnie jak w metodzie 1, budowa toru powinna odzwierciedlać normalne warunki typowego toru z uwzględnieniem profilu szyny, szerokości toru i stanu utrzymania. Jazdy pomiarowe należy tak zaplanować, aby koło o najmniejszym pionowym nacisku było kołem prowadzącym. W metodzie tej należy zniwelować powstawanie siły wzdłużnej na skład pociągu, a badany pojazd nie może być hamowany. Badania należy przeprowadzić minimum 3 razy z prędkością nie przekraczającą 10 km/h. Wielkościami mierzonymi podczas tego testu na łuku pomiarowy są: siły prowadzące na wewnętrznym i zewnętrznym kole badanego pojazdu Y_i, Y_a, siła pionowego nacisku na wewnętrznym kole badanego pojazdu Q_i oraz kąt nabiegania zestawu prowadzącego α. Powyższe parametry mogą być mierzone przez urządzenia umieszczone w torze lub na pojeździe. W przypadku, gdy urządzenia pomiarowe umieszczone są w torze to ich lokalizacja znajduje się w dwóch strefach. Pierwsza strefa znajduje się na początku łuku w odległości powyżej 3 m aż do $2a^*$ ($2a^*$ - odległość środków wózków lub odległość osi w pojazdach bezwózkowych). Lokalizacja tej strefy zapewnia pomiar sił podczas obrotu wózka względem pudła, co jest bardzo istotne przy takich konstrukcjach pojazdów. W strefie tej przewidziano minimum 3 punkty pomiarowe. Kolejna strefa pomiarowa powinna być tak usytuowana, tak aby cały pojazd znajdował się w łuku. Początek strefy należy umieścić w odległości powyżej $2a^+ + 2a^*$ ($2a^+$ - rozstaw skrajnych zestawów kołowych w wózku) licząc od początku łuku. Również strefa ta powinna zawierać minimum 3 punkty pomiarowe.

Podczas realizacji tej metody badań siły prowadzące Y_i i Y_a należy rejestrować dla każdej pozycji pomiarowej. Ich ocena dokonywana jest przy pomocy wartości średniej $Y_{i,med}$ i $Y_{a,med}$ z punktów pomiarowych oddzielnie dla każdej ze stref pomiarowych. Kierunek siły Y_i w większości przypadków jest przeciwny do siły Y_a. Mając na uwadze mierzone parametry, takie jak sił (Y_i, Q_i) na szynie wewnętrznej i kąt nabiegania zestawu prowadzącego α zakłada się, że współczynnik tarcia między kołem a szyną jest bliski wartości granicznej tarcia koła o szynę. Zatem tor przed badaniami należy przygotować podobnie, jak to opisano w metodzie 1.

Ocenę spełnienia wymagań bezpieczeństwa, a tym samym minimalizacji ryzyka wykolejenia, podczas jazdy po zwichrowanym torze należy wykonać dla każdego zestawu

kołowego. W tym przypadku, wg normy EN 14363:2016 [11], wykorzystujemy zależności (5) i (6) określające wartość współczynnika wykolejenia oraz jego limit.

$$\left(\frac{Y}{Q}\right)_{ja} = \frac{Y_{ja,med}}{Q_{jk,\min} + \Delta Q_{jH}}, \tag{5}$$

gdzie: $Y_{ja,med}$ oznacza quasistatyczną siłę prowadzącą określoną na podstawie przejazdu pojazdu przez łuk o promieniu R = 150 m, $Q_{jk,min}$ jest minimalną siłą pionowego nacisku obliczoną na podstawie próby wichrowania, ΔQ_{jH} reprezentuje zmianę siły pionowego nacisku koła spowodowaną momentem sił prowadzących. Kryterium oceny ryzyka wykolejenia w tym przypadku przyjmuje postać formuły (6).

$$(Y/Q)_{ja} \leq (Y/Q)_{lim}, \tag{6}$$

zgodnie z podejściem prezentowanym przez Nadal'a, dla koła o kącie pochylenia obrzeża 70° i współczynniku tarcia koła o szynę μ = 0,36. Granicza wartość współczynnika wykolejenia wynosi $(Y/Q)_{lim}$ = 1,2.

Metoda 3, zaczerpnięta z normy [11], dotyczy testów pojazdu na stanowisku do wichrowania i stanowisku do pomiaru momentu oporowego wózka względem nadwozia. Może być ona użyta do badania pojazdów zbudowanych w konwencjonalnej technologii, czyli pojazdy które są eksploatowane w normalnych warunkach i odpowiadają całkowicie lub ich części konstrukcyjne związane z zachowaniem własności biegowych odpowiada ustalonemu poziomowi wiedzy. W tym przypadku są to pojazdy z wózkami dwuosiowymi, dwa wózki na pojazd oraz przy kącie pochylenia obrzeża koła γ w zakresie 68° a 70°. W związku z powyższym, metody tej nie można zastosować do pojazdów z wspólnym wózkiem oraz do pojazdów szynowych wyposażonych w wózki trzyosiowe.

Do oceny bezpieczeństwa jazdy po zwichrowanym torze oraz ryzyka wystąpienia wykolejenia, w oparciu o metodę 3, należy użyć dwóch stanowisk badawczych, tzn. stanowiska do pomiaru nacisków kół i stanowiska do pomiaru momentu oporowego wózka względem nadwozia. Stanowisko używane do pomiarów nacisków kół, podobnie jak w metodzie 2, powinno być wyposażone w urządzenia służące do unoszenia i opuszczania kół. Niezależne przemieszczenie kół musi być realizowane przynajmniej na dwóch zestawach kołowych jednego wózka. Wymagana w badaniach wichrowatość toru jest wyznaczana z takich samych zależności jak w metodzie 1. Zatem zakres przemieszczania kół podczas pomiarów nacisków kół w metodzie 3 jest większy od zakresu wyznaczonego w metodzie 2. Algorytm przemieszczania kół pojazdu jest taki sam jak w metodzie 2. W trakcie wichrowania mierzone jest przemieszczenie Δz_{jk} kół oraz siły nacisku Q_{jk} wszystkich kół w sposób ciągły. Na podstawie tych danych wyznaczony zostaje pionowy nacisk badanego koła Q_0 na poziomym torze oraz spadek nacisku koła ΔQ wywołany maksymalną wichrowatością.

Drugi etap badań związany z pomiarem momentu oporowego wózka względem nadwozia realizowany jest na stanowisku stacjonarnym. Takie stanowisko umożliwia obrót wózka w lewo i w prawo względem nadwozia o zadany kąt. Wymagana prędkość obrotu stanowiska jest stała i wynosić 1°/s w zakresie 75% kąta skręcenia wózka względem nadwozia. Ze względu na obciążenie wagonu ładunkiem podczas eksploatacji, pomiary momentu oporowego wózków względem nadwozia powinny być przeprowadzone dla pojazdu w stanie próżnym i ładownym. Podczas badań wózek powinien być połączony z nadwoziem pojazdu za pomocą wszystkich przewidywanych połączeń. Istotą tej metody jest wyznaczenie stosunku odciążenia koła prowadzącego do siły pionowego nacisku przy braku wichrowatości oraz współczynnika X, charakteryzującego zachowanie się wózka na łukach o małym promieniu.

Dopuszczalne wartości ww. parametrów opisują zależności (7) i (8). Zgodnie z normą EN 14363:2016 [11] pojazd uznaje się za bezpieczny jeżeli spełnia jednocześnie dwa kryteria.

$$\frac{\Delta Q}{Q_0} \leq 0,6\,, \tag{7}$$

gdzie: ΔQ jest odchyłką od Q_0 w warunkach maksymalnej wichrowatości, Q_0 określa średnią siłą pionowego nacisku koła badanego zestawu na poziomym torze oraz wskaźnik X dla wagonów towarowych zależy od obciążenia osi. Współczynnik X należy wyznaczyć z zależności (8).

$$X = \frac{M_{z,Rmin}}{2a^{+}2Q_0} \quad , \tag{8}$$

gdzie: $M_{z,R\,min}$ – wartość momentu oporowego wózka względem nadwozia dla kąta $\psi = a^*/R_{min}$, $2a^+$ – rozstaw zestawów kołowych w wózku (baza wózka Y25 wynosi 1,8 m), $2Q_0$ – nacisk badanego zestaw kołowego. W przypadku wagonów towarowych wartość kryterialną współczynnika X wyznaczono z wykresu prezentującego graniczną wartość współczynnika X w zależności od nacisku osi na tor $2Q_0$ (Rys. 5).

3. Badania eksperymentalne

3.1 Obiekt badań

Do badań eksperymentalnych wykorzystano typowy wagon węglarkę serii Eanoss wyposażony w dwa standardowe wózki rodziny Y25. Badany pojazd jest przeznaczony do przewozu kruszywa, węgla i materiałów sypkich. Jego eksploatacja może odbywać się na torach o szerokości 1,435 m. Parametry geometryczne wagonu to: całkowita długość ze zderzakami LUP = 14,04 m, maksymalna szerokość 3,038 m, natomiast maksymalna wysokość to 3,43 m. Masa badanego wagonu wynosiła 20,3 t. Wagon został zaprojektowany do maksymalnego obciążenia zestawu kołowego 22,5 t, czyli do masy brutto wynoszącej 90 t. W stanie próżnym eksploatacja wagonu może odbywać się z maksymalną prędkości 120 km/h, natomiast w stanie ładownym do 100 km/h.

Parametry niezbędne do dowodu, że wagon może być bezpiecznie eksploatowany to: rozstaw czopów skrętu $2a^*$ = 9,0 m, rozstaw zestawów kołowych w wózku $2a^+$ = 1,8 m.

3.2 Opis metody badań

W celu sprawdzenia czy badany wagon węglarka serii Eanoss może być bezpiecznie eksploatowany wykorzystano opisaną w normie EN 14363:2016 metodę nr 3.

Podczas badań eksperymentalnych wagonu towarowego użyto dwóch stanowisk stacjonarnych. W pierwszym etapie badań było to stanowisko do pomiarów nacisków kół, natomiast w drugim stanowisko do pomiaru momentu oporowego wózków względem nadwozia.

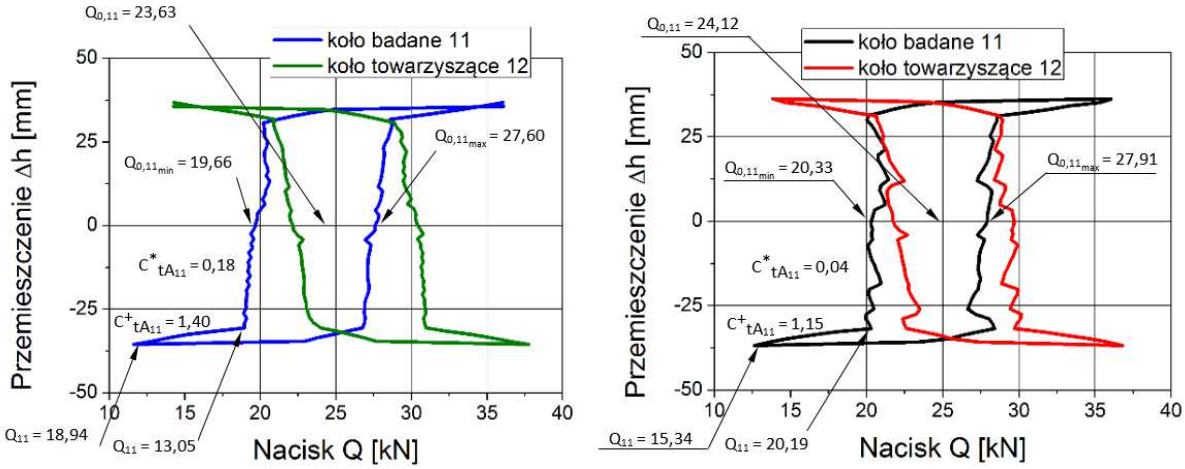
Pomiary nacisków kół przeprowadzono na stanowisku TENSAN-PLW, które charakteryzuje się możliwością niezależnego przemieszczania poszczególnych kół. Na omawianym stanowisku Instytut Kolejnictwa posiada własne oprogramowanie, za pomocą którego podczas badań realizowane jest pionowe wymuszenie poszczególnych kół, w czasie którego prowadzona jest rejestracja ich nacisków oraz przemieszczenia pionowego. Użyte do badań oprogramowanie zostało napisane w oparciu o wytyczne do wichrowania pojazdów podane w normie EN 14363:2016.

Stanowisko TENSAN-PLW charakteryzuje się specjalistycznymi modułami pomiarowymi, w skład których wchodzą ułożyskowane w osi toru wahacze widoczne na rysunku 2. Ich przemieszczania, względem zerowego poziomu toru, wymuszone są przez siłowniki hydraulicznych. Lokalizacja modułów pomiarowych oraz ich długość umożliwia

ustawienie na nich każdego koła pojazdu oddzielnie, a tym samym pozwala na indywidualny pomiar nacisku każdego z nich (Rys. 2). Całkowita długość stanowiska wynosi 22,22 m. Zakres maksymalnego pionowego nacisku na pojedynczy wahacz pomiarowy wynosi Q = 200 kN a przemieszczenie wahacza mieści się w zakresie Δh = ± 0,07 m. W przypadku pomiarów nacisków kół pojazdów przegubowych, których długość jest większa od długości stanowiska, tor przed stanowiskiem i za stanowiskiem jest zniwelowany, co pozwala uzyskać pionową różnicę położenia szyn.

3.3 Pomiary i otrzymane wyniki

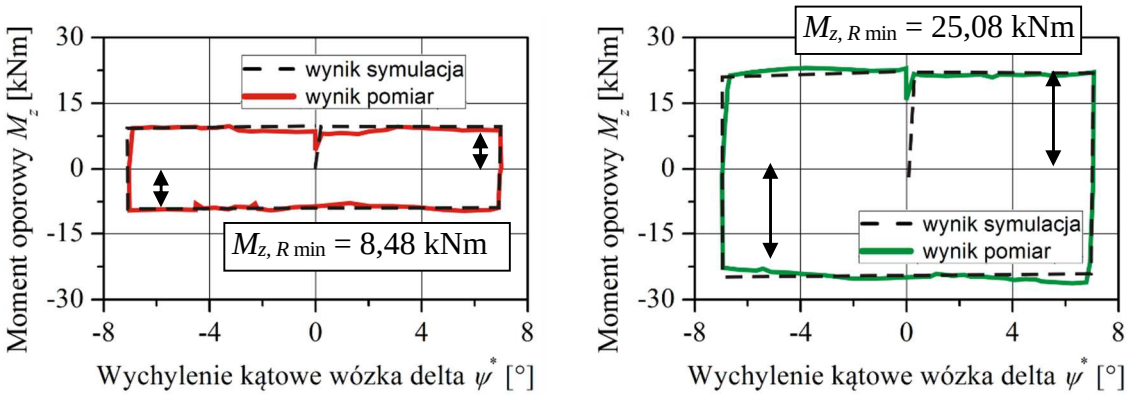
Podczas badań rozpatrywanego wagonu węglarki serii Eanoss naciski kół mierzono w trakcie przemieszczania poszczególnych kół w górę i w dół z neutralnego poziomu toru. Przeprowadzono dwie próby wichrowania, tj. badanie na nieruchomym zestawie kołowym (zestaw badany, który podlega ocenie nie ulega przemieszczeniu podczas badań, przemieszczane są koła w drugim wózku) i badanie na ruchomym zestawie kołowym (zestaw badany jest przemieszczany). W pracy w celu ograniczenia liczby jego stron zamieszczono tylko wyniki analizy sił koła pierwszego prawego wózka przedniego - wykresy na rysunku 3. Na ich podstawie określono minimalny i maksymalny nacisk koła ($Q_{0,jkx\ min}$ i $Q_{0,jkx\ max}$), występujący na poziomym torze podczas próby wichrowania. Ze średniej arytmetycznej otrzymanej z tych nacisków wyznaczono nominalny nacisk koła z wyeliminowaniem wpływu histerezy tarcia, który w analizowanym przypadku wyniósł $Q_{0,jkx}$=24,12 kN. Badania te pozwoliły także określić sztywności skrętne badanego pojazdu mierzone na bazie rozstawu czopów skrętu $C^{*}_{tA_{jk}}$ i osi w wózku $C^{+}_{tA_{jk}}$. Powyższe parametry posłużyły do wyznaczenia minimalnego nacisku koła podczas przejazdu pojazdu przez tor o danej wichrowatości. W dalszym etapie badań zostały one użyte w badaniach teoretycznych poprzez implementacje ich w modelu numerycznym pojazdu towarowego prezentowanym w rozdziale 4 tego artykułu.



Rys. 3. Pomiary siły pionowej na stanowisku, wariant pomiar na nieruchomym (lewy) i ruchomym zestawie kołowym (prawy)

W kolejnym etapie badań opisanego powyżej wagonu było stanowisko do pomiaru momentu oporowego wózka względem nadwozia należące do Instytutu Kolejnictwa. Podczas testów maksymalny kąt obrotów stanowiska wynosił ψ= ± 10°, a prędkość obrotu stanowiska nastawiana była na trzy wartości tj. 0,2, 0,6, 1,0 [1°/s]. Finalnie, zgodnie z normą [11], do oceny brano pod uwagę wyniki zarejestrowane przy prędkości obrotu stanowiska równej 1,0 [1°/s]. Pozostałe pomiary, nie podlegające ocenie, służą jedynie do uzyskania odpowiedniej współpracy elementów ciernych połączenia właściwego pomiaru wózka z nadwoziem wagonu. Badaniu poddano wagon w stanie próżnym, jak i ładownym tj. z obciążeniem brutto 89,7 t.

Przy każdej prędkości obrotu wózka względem nadwozia i dla każdego obciążenia wagonu przeprowadzono po jednym pomiarze momentu oporowego wózka względem nadwozia. Otrzymane podczas badań eksperymentalnych i symulacyjnych wyniki momentu oporowego na czopie skrętu, przy dwóch wariatach obciążenia wagonu i prędkości obrotu 1,0 [1°/s], pokazano na wykresach (Rys. 4).

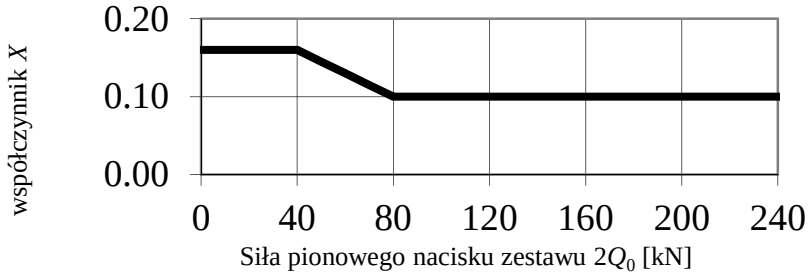


Rys. 4. Pomiar momentu oporowego wózka względem nadwozia wagonu towarowego, wagon próżny (lewy wykres), wagon ładowny (prawy wykres) prędkości pomiaru $V_{obr.}$ = 1,0 [˚/s]

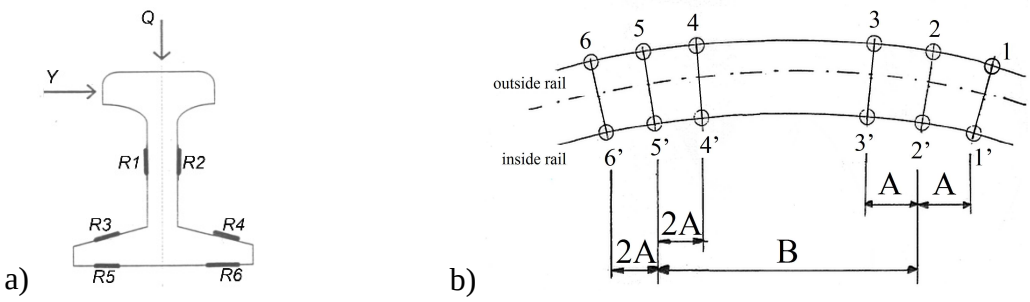
Wartości liczbowe momentu oporowego przy kilku prędkościach obrotowych oraz różnym obciążeniu pojazdu zamieszczono w tablicy 1. Prezentowane wyniki otrzymano przy wykorzystaniu kryteriom oceny wg granicznej wartości wskaźnika X (Rys. 5). W obu przypadkach obciążenia występuje wzrost momentu oporowego wraz ze zwiększeniem prędkości podczas obrotu. Wskazuje to na istotny wpływ tarcia o charakterystyce degresywnej między wózkiem a nadwoziem wagonu. Model numeryczny wagonu towarowego wykorzystany do badań symulacyjnych zamieszczonych na rysunkach 4 opisano w dalszej części pracy.

Tablica 1. Wyniki pomiarów momentu oporowego wagonu towarowego

typ wózka	stan wagonu	wartość momentu oporowego $M_{z,\,R\,\min}$ [kNm] przy różnych prędkościach pomiaru		
		$V_{obr.}$ = 0,2 [˚/s]	$V_{obr.}$ = 0,6 [˚/s]	$V_{obr.}$ = 1,0 [˚/s]
Y25	próżny	9,54	8,72	8,48
	ładownym	26,07	25,94	25,08



Rys. 5. graniczna wartość współczynnika X dla pojazdów towarowych w zależności od nacisku osi na tor



Rys. 6. Lokalizacja układów tensometrycznych na szynie w jednym punkcie pomiarowym (a) oraz rozmieszczenie wszystkich przekrojów pomiarowych 1, 2, 3, 4, 5 i 6 na łuku toru testowego (b)

Wymienione powyżej stanowisko do dynamicznych badań wartości sił w strefie kontaktu koło-szyna stanowił tor badawczy ułożony w łuku o promieniu R = 150 m na długości 95 m. W torze tym zastosowano szyny typu S-49 o pochyleniu 1:40, zamocowane sprężyście do podkładów strunobetonowych INBK-7 przez system przytwierdzenia SB-3. Rozstaw podkładów wynosi 0,60 m, a w łuku zastosowano nominalne poszerzenie toru o wartości e = 0,005 m. Pomiar sił odbywa się na poziomie szyny przy użyciu stałoprądowych podwójnych mostków tensometrycznych (Rys. 6a) naklejonych na obie szyny, zlokalizowanych w przekrojach pomiarowych 1, 2, 3, 4, 5 i 6 (Rys. 6b). Rozstaw przekrojów pomiarowych jest wyznaczony w oparciu o normę EN 14363 [11]. Wartość siły pionowej Q określamy z liniowej relacji wartości sygnałów z układu pomiarowego (mostki tensometryczne R_1 i R_2 na szyjce szyny), które są poddane wzmocnieniu przez dedykowane im wzmacniacze pomiarowe. Liniowa relacja sygnału mierzonego odpowiadającego za siłę poziomą Y, skorelowaną z oddziaływaniem siły pionowej Q działającej na pkt. pomiarowy, wyznaczono na stanowisku kalibracyjnym i zapisano je zależnością (9).

$$\begin{cases} Y_0 = Y - K_0 Q \\ Y_u = Y - K_u Q \end{cases}, \tag{9}$$

gdzie: Y_0 oznacza pomiar z podwójnego mostka tensometrycznego naklejonego na górną część stopki szyny (mostki R_3 i R_4), Y_u oznacza mostek z dolnej części stopki (mostki R_5 i R_6), Q opisuje siłę pionową oddziaływania koła na szynę, a K_0 i K_u są współczynnikami wzmocnienia wzmacniaczy pomiarowych. Z powyższego układu równań otrzymujemy równanie (10) określające wartość siły bocznej Y w punkcie pomiarowym.

$$Y = \frac{Y_0}{(1 - \frac{K_0}{K_u})} - \frac{Y_u}{\left(\frac{K_u}{K_0} - 1\right)} . \tag{10}$$

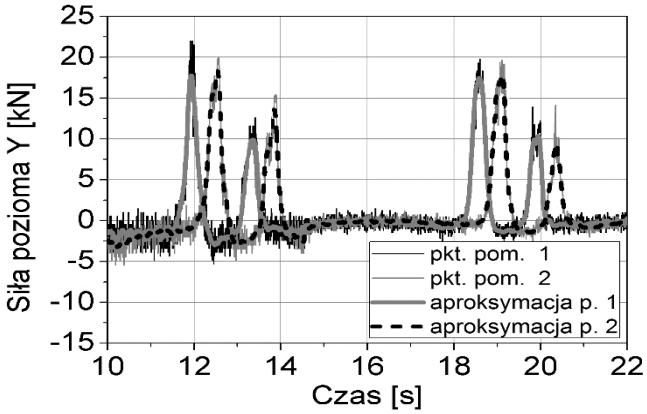
Eksperymentalne badanie współczynnika Y/Q przeprowadzono podczas przetaczania wagonu towarowego po torze z prędkością v=5 km/h. Przebiegi czasowe wartości sił Y i Q zmierzone w przekrojach 1, 2, 4, 5 na szynie zewnętrznej łuku toru i określone z zależności 9 i 10 przedstawiają wykresy na rysunkach 7 ÷ 9. Na podstawie otrzymanych danych z każdego punktu pomiarowego wyznaczono maksymalną wartości współczynników wykolejenia dla każdego koła pojazdu szynowego, które zamieszczono w tablicy 2. W tabeli tej poszczególne koła badanego pojazdu oznaczono symbolem W z indeksami i, j. Indeks i oznacza nr osi pojazdu, a indeks j stronę zestawu kołowego (j=1 strona prawa, j=2 strona lewa).

Na podstawie analizy wyników zarejestrowanych w danych przekrojach pomiarowych podczas testowych przejazdów zaobserwowano, że zestawy kołowe wózków pojazdu

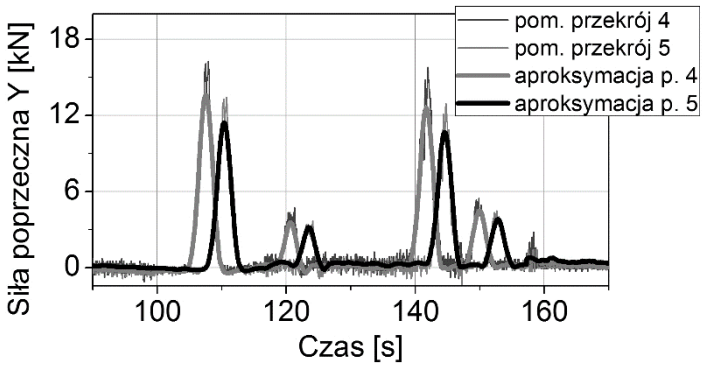
chwilowo odciążały poprzecznie tor. Wywołane to zostało różnymi kątami nabiegania między pierwszym i drugim zestawem kołowym danego wózka. W analizowanym przypadku podczas wjazdu w łuk toru zaobserwowano różnice sił poprzecznych zestawów pojedynczego wózka na poziomie 50% i 36%, odpowiednio w wózku pierwszym i drugim badanego wagonu (Rys. 7). Podczas wyjazdu z łuku toru różnice sił poprzecznych miedzy zestawami pojedynczego wózka są znacznie większe i wynoszą ponad 60% (Rys. 8). Takie zachowanie zestawów kołowych przekładało się także na zmniejszenie pionowych nacisków drugich zestawów kołowych na tor, które w analizowanym przypadku wynosiło w przedziale 7-15% w stosunku do nacisków zestawu pierwszego (Rys. 9). Takie zachowanie znacząco wpływa na wzrost wartości współczynnika wykolejenia. Podobne obserwację otrzymano z analizy wyników numerycznych wykonywanych powstałym modele analizowanego pojazdu opisanym w kolejnej sekcji artykułu. Zjawiska dominującego oddziaływania poprzecznego atakującego zestawu kołowego (pierwszego zestawu w wózku) podczas jazdy w łuku toru występuje często w pojazdach szynowych [33, 43]. W takich warunkach przeważnie wózki wężykują, a wtedy przy największej amplitudzie wężykowania podłużna oś symetrii wózka jest odchylona od podłużnej osi symetrii toru o pewien kąt α, zwany kątem natarcia [3, 6, 30]. W takim przypadku wzrasta także zużycie zestawów kołowych oraz powierzchni tocznej szyn w postaci poligonizacji i korugacji [2].

Tablica 2. Wartości współczynnika wykolejenia określone z badań eksperymentalnych na torze

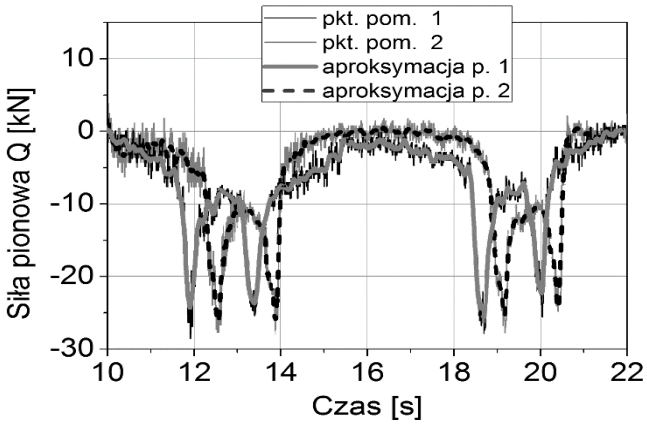
Koło W_{ij}	**11**	**12**	**21**	**22**	**31**	**32**	**41**	**42**
Y/Q [kN]	0.73	0.72	0.21	0.24	0.59	0.69	0.66	0.62



Rys. 7. Przebiegi sił poprzecznych Y w dwóch przekrojach pomiarowych (2 i 3) podczas przetaczania wagonu towarowego po torze testowym



Rys. 8. Przebiegi sił poprzecznych Y w dwóch przekrojach pomiarowych (4 i 5) podczas przetaczania wagonu towarowego po torze testowy



Rys. 9. Siły pionowe Q w dwóch przekrojach pomiarowych (4 i 5) podczas przetaczania wagonu towarowego po torze testowym

4. Analiza numeryczna zjawisk towarzyszących ryzyku wykolejenia

4.1 Model numeryczny pojazdu szynowego

Aby określić teoretycznie wartości sił w strefach kontaktu kół pojazdu z szynami toru powstał model wagonu posadowionego na wózkach serii Y25. Model fizyczny rozpatrywanego pojazdu potraktowano jako układ brył sztywnych połączonych ze sobą za pomocą elementów sprężysto-tłumiących (Rys. 10a i 10b). Takie podejście w modelowaniu nazywa się metodą układów wieloczłonowych [13,44] i jest bardzo często stosowane przez badaczy we własnych kodach do analizy dynamiki pojazdów szynowych [3,26,45] oraz w programach komercyjnych, tj. Vampire, ViGrade/VI-Rail, Autodyn, Simpack, UM Loco. W metodzie tej elementy konstrukcyjne pojazdu traktowane są jako nieodkształcalne ciała, a elementy zawieszenia opisują elementy podatne [4,8]. Ograniczenia w ruchu tych ciał wynikają z narzuconych więzów holonomicznych całkowalnych [37].

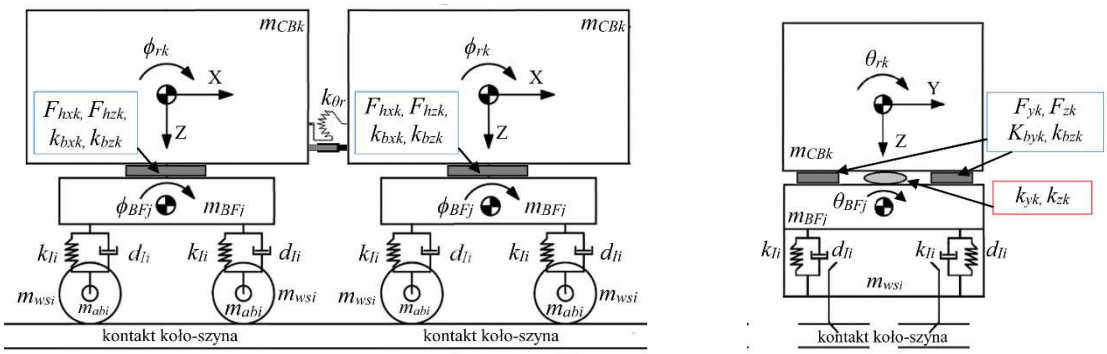
Rozpatrywany w pracy model wagonu podzielono na trzy bazowe elementy, którymi są zestawy kołowe, ramy wózków oraz człon pojazdu (nadwozie). Człon podzielono na dwie oddzielne bryły, aby uwzględniać w modelu pojazdu sztywność skrętną nadwozia $K\theta_r$ wyznaczoną eksperymentalnie (Rys. 4). Zestaw kołowy opisano jako bryłę o trzech stopniach swobody, gdzie przemieszczenia poprzeczne oznaczono symbolem (y_i), kąt natarcia (φ_i) oraz galopowanie (ϕ_i). Użyte w opisie symboli indeks i odpowiada poszczególnym zestawom kołowym, przyjmując wartości i=1, 2, 3, 4. Ramę wózka reprezentuje bryła o pięciu stopniach swobody, które odpowiadają przemieszczeniu poprzecznemu (y_{rj}), przemieszczeniu pionowemu (z_{rj}), wężykowaniu (φ_{rj}), kątowi kołysania (θ_{rj}) oraz kątowi galopowania (ϕ_{rj}). W tym przypadku indeks wynosi j=1,2, gdyż występują dwa wózki. W przypadku nadwozia rozpatrywanego pojazdu, bryły je opisujące mają pięć stopni swobody. Należą do nich przemieszczenie poprzeczne (y_{nk}), przemieszczenie pionowe (z_{nk}), kąt wężykowania (φ_{nk}), kąt kołysania (θ_{nk}), kąt galopowania (ϕ_{nk}), gdzie k=1,2, gdyż występuje podział nadwozia na dwie bryły. Podsumowując, stworzony do badań numeryczny model pojazdu szynowego miał 64 stopnie swobody.

W rozpatrywanym przypadku model matematyczny wagonu towarowego opisano przez układ równań różniczkowych, które wyprowadzono korzystając z równań Lagrange'a II rodzaju. W tym podejściu współrzędne uogólnione przyjmują postać przemieszczeń linowych lub kątów obrotu. Ruch takiego pojazdu opisany jest równaniami różniczkowalnymi zwyczajnymi drugiego rzędu. Przy założeniu, że oscylacje poszczególnych brył modelu

względem układu odniesienia są niewielkie, układ taki można zapisać w postaci zlinearyzowanego układu równań (11) zapisanego w postaci macierzowej [14].

$$[\mathbf{M}\, d^2/dt^2 + \mathbf{C}\, d/dt + \mathbf{K}]\cdot \mathbf{q} = \mathbf{F} \tag{11}$$

gdzie: $\mathbf{q}=\{ y_i, y_{rj}, y_n, z_{rj}, z_n, \varphi_i, \phi_{rj}, \phi_n, \phi_i, \theta_{rj}, \theta_{n,}\}^T$ – wektor współrzędnych uogólnionych układu, **M** – symetryczna macierz bezwładności, **C** - macierz tłumienia, **K** - macierz sztywności, **F** - wektor sił, d/d*t* – operator różniczkowy. Poprzez zastosowanie metod Newtona-Raphsona, powstały układ liniowych równań algebraicznych, który rozwiązywany jest w każdej iteracji z krokiem czasowym Δt=0.001 s.



a) b)

Rys. 10. Fizyczny model analizowanego pojazdu (a) boczny widok, (b) czołowy widok pojazdu

Tabela 3. Parametry bezwładnościowe modelu wagonu towarowego

bryły	masa [kg]	I_{xx} [kg·m^2]	I_{yy} [kg·m^2]	I_{zz} [kg·m^2]
pudło m_{CB}	16000,00	47500,00	51000,00	50050,00
wózek m_{BF}	2000,00	1975,00	2850,00	1560,00
obud. łożyskowa m_{ab}	25,00	10,00	10,00	10,00
zestaw kołowy m_{ws}	1300,00	688,00	100,00	688,00

Szczegółowe parametry bezwładnościowe elementów modelu zestawiono w tablicy 3. Rozstaw kół zestawu kołowego wynosił 1,5 m, promień toczny kół *R*= 0,42 m, rozstaw wózków 9 m, a rozstaw zestawów kołowych 1,8 m, rozstaw ślizgowych łożysk bocznych 1,7 m. Sztywności sprężyny zawieszenia pierwszego stopnia wynoszą odpowiednio $k_{Ix}=k_{Iy}$=4000 kNm oraz k_{Iz}=3950 kNm, sztywności czopu skrętu przyjęto na poziomie $k_x=k_y= k_z$ =10 MNm, sztywność poprzeczna i pionowa ślizgowych łożysk bocznych wynosi odpowiednio k_{sl_z}=350 kNm i k_{sl_z}=500 kNm, parametry łożysk zaczerpnięto z pracy [36]. Czop skrętu opisano jako połączenie sferyczne o trzech rotacyjnych stopniach swobody, w którym zmianę sił tarcia wyznaczono z zależności

$$\begin{aligned} F_x &= (W_x \,.\, \chi) \,/\, \sqrt{(1 + \left(\frac{Wm\,\chi s}{rN\mu_2}\right) \text{^}2)}, \\ F_y &= (W_y \,.\, \chi) \,/\, \sqrt{(1 + \left(\frac{Wm\,\chi s}{rN\mu_2}\right) \text{^}2)}, \\ F_z &= (W_z \,.\, \chi) \,/\, \sqrt{(1 + \left(\frac{Wm\,\chi s}{rN\mu_2}\right) \text{^}2)}, \end{aligned} \tag{12}$$

gdzie: W_x, W_y, W_z opisują względną prędkość obrotową czopu wokół osi X, Y, Z; W_m jest bezwzględną prędkością między pudłem wagonu a wózkiem w centralnej punkcie

czopu, χ_s oznacza styczną w punkcie początkowym funkcji przenoszenia siły/prędkości o wartości $3{,}0 \cdot 10^6$ Ns/m, N opisuje normalną siłę skierowaną w kierunku Z, μ_2 określa współczynnik tarcia o wartości 0,19, r jest to promień środkowej krzywizny czopu o wartości 0,19 m [36].

Siły tarcia F_{bx}, F_{bx} w płaszczyznach łożysk bocznych, poprzez które wsparto dodatkowo nawozie na wózku określono z zależności (13) [5]

$$\begin{aligned} F_{bx} &= (V_x \,.\, \chi) \,/\, \sqrt{\left(1 + \left(\frac{Wm\,\chi l}{N\mu_1}\right)\text{^}2\right)} \\ F_{by} &= (V_y \,.\, \chi) \,/\, \sqrt{\left(1 + \left(\frac{Wm\,\chi l}{N\mu_1}\right)\text{^}2\right)} \end{aligned} \quad , \tag{13}$$

gdzie: V_x, V_y określają względne prędkości w płaszczyźnie łożyska w kierunkach X, Y, W_m oznacza całkowitą względna prędkość na płaszczyźnie X-Y, χ_l opisuje styczną w punkcie początkowym funkcji przenoszenia siły/prędkości o wartości $3{,}0 \cdot 10^6$ Ns/m, N oznacza siłę wywołaną ciężarem pudła wagonu działająca w kierunku normalnym do płaszczyzny XY, μ_1 jest współczynnikiem tarcia równym 0,36 [5]. Ze względu na fakt, że praca dotyczy analizy wykolejenia pojazdu szynowego, strukturę modelu wagonu towarowego oraz jego zapis matematyczny opisano w pracy ogólnie zależnością macierzową (11). Natomiast, więcej uwagi poświęcono na opis kontaktu koło-szyna i metodologię wyznaczenia sił w strefie kontaktu koło szyna oraz współczynnika wykolejenia wynikającego z relacji między tymi siłami.

4.2 Model kontaktu koło-szyna

Model matematyczny wagonu towarowego posadowionego na wózkach serii Y25 zintegrowano z algorytmami i procedurami numerycznymi wyznaczającymi kontakt kół z szynami. Procedury numeryczne dotyczące kontaktu koło-szyna posłużyły do wyznaczenia wartości sił oraz obszarów ich działania w strefach kontaktu. Model kontaktu oparto na uproszczonej teorii Kalkera [22] i algorytmie FASTSIM [23]. Aby obliczyć styczne siły styku wyznaczono normalne siły nacisku, współczynnik tarcia przyjęto na poziomie 0,36 [14,26 35], długość półosi a i b elipsy pola obszaru kontaktu (Rys. 11a) obliczono przy użyciu teorii Hertza [19]. Wartości pełzania uwzględniono w postaci względnego sztywnego poślizgu. Relacje między tymi parametrami opisano zależnością (14).

$$\begin{bmatrix} rs_x \\ rs_y \\ rs_z \end{bmatrix} = \frac{1}{vu_x} \begin{bmatrix} sv_x \\ sv_y \\ sv_y \cdot \sin(|\alpha|) + sv_z \cdot \cos(|\alpha|) \end{bmatrix}, \tag{14}$$

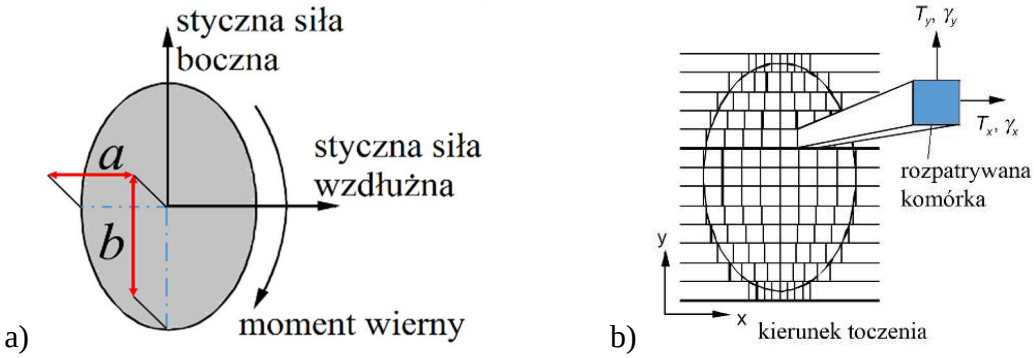
gdzie: rs_x, rs_y określają względny sztywny poślizg/pełzanie w kierunku podłużnym X i bocznym Y, rs_z oznacza spin, ω opisuje kąt kontaktu, vu jest prędkością ruchomego układu odniesienia, której wartość jest równa prędkości pojazdu, sv określa prędkość poślizgu. Prędkości poślizgu sv rzutowane na dany kierunek wyznaczono z zależności (15).

$$\begin{bmatrix} sv_x \\ sv_y \\ sv_z \end{bmatrix} = \begin{bmatrix} vu_x \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ vr_y \\ 0 \end{bmatrix} + \bar{\omega}_w \times \bar{r}, \tag{15}$$

gdzie: $\bar{\omega}_w$ oznacza względna prędkość kątowa koła, $\bar{r}$ jest współrzędną punktu kontaktowego w układzie odniesienia połączonej z centrum masy koła, vr opisuje prędkość względna środka masy koła określona w ruchomym układzie odniesienia.

W kolejnym kroku wyznaczono styczne siły kontaktowe T_x i T_y stosując procedurę FASTSIM [23]. Algorytm tej procedury dokonuje podziału eliptycznej strefy kontaktu, na mniejsze obszary/komórki. W każdej z tych komórek wyznaczane są naprężenia styczne oraz mikropoślizgi w dwóch kierunkach, wzdłużnym v_x i poprzecznym v_y do kierunku jazdy pojazdu (Rys. 11b). Następnie wyznaczamy strefy przylegania w obszarze kontaktu przy $\gamma = 0$. Parametry geometrii kontaktu jako wielkości wejściowe do procedury FASTSIM zostały stablicowane w zależności od przemieszczenia poprzecznego poszczególnego koła (Rys. 12 i 13), przez co zredukowano czas obliczeniowy podczas symulacji dynamiki ruchu pojazdu.

W dalszym etapie procedury obliczeniowej wyznaczane i obliczane są prawe stron dynamicznych równań ruchu (11) dla odpowiednich elementów pojazdu oraz przygotowanie ich do dalszego rozwiązania poprzez całkowanie numeryczne tych równań ruchu opisujących rozpatrywany układ pojazd szynowy-tor. Opisane w tym rozdziale podejście pozwoliło wyznaczyć wartości sił w strefach kontaktu koło-szyna, które w dalszym etapie posłużyły do obliczenia wartości współczynnika wykolejenia *Y*/*Q*.

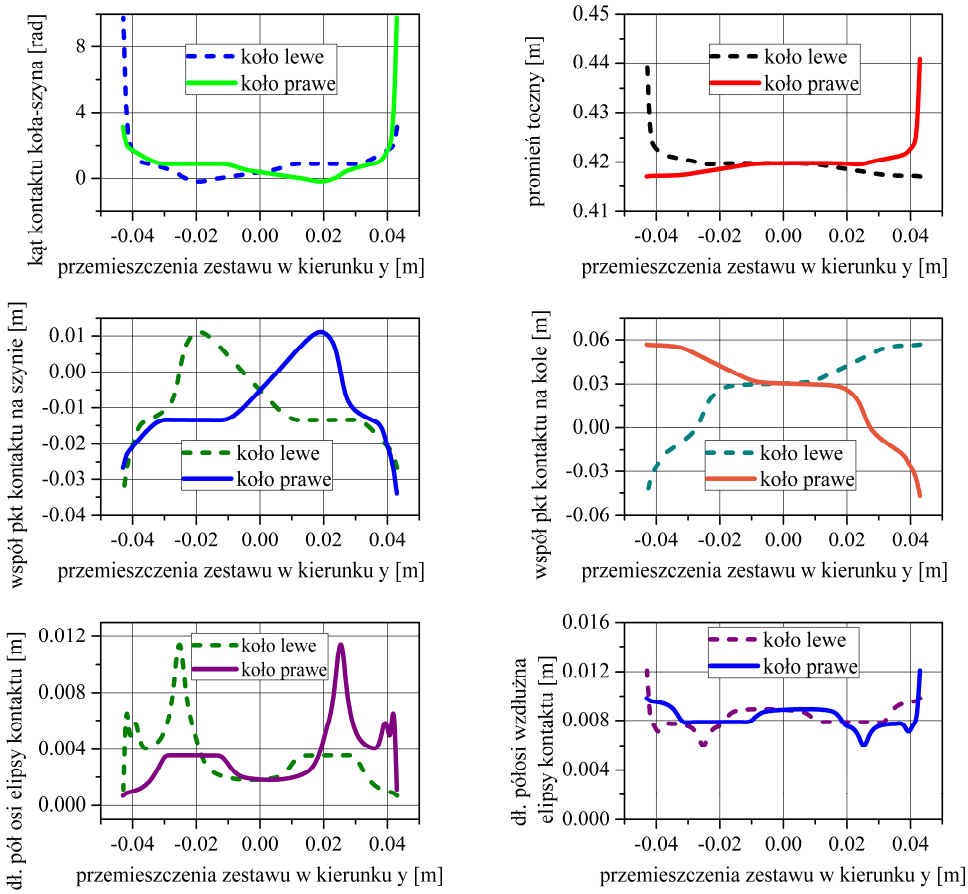


Rys. 11. Siły styczne w strefie kontaktu koło-szyna, półosie *a* i *b* elipsy kontaktu oraz podział jej powierzchni na elementy

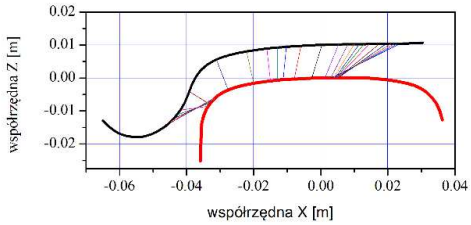
W modelu symulacyjnym do badań dynamiki wagonu towarowego wykorzystano rzeczywiste nominalne zarysy kół kolejowych UIC 60 i szyny 49E1 [38], w konfiguracji pochylenia szyn w torze 1:40 i przedstawionym na rysunku 14.

4.3 Wyniki badań numerycznych

Zakres przeprowadzonych analiz przy wykorzystaniu powyżej opisanego modelu wagon towarowy-tor dotyczy dynamiki ruchu pojazdu po torze zakrzywionym typu S o promieniu *R*=150 m i wichrowatości 3‰ (Rys. 12). Jako oryginalny wkład pracy do badań przyjęto wichrowatość toru pozwalająca podczas pojedynczego przejazdu wyznaczyć badania dla pojazdu i wózka, poprzez wprowadzenie dodatkowego wzniosu w pionowym profilu toru (Rys. 12). Podczas tych badań przyjęto najbardziej niekorzystny wariant konfiguracji eksploatacyjnej pojazdu, czyli pojazd bez obciążenia. Podczas przejazdu rejestrowane będą siły pionowa *Q* i boczne (poprzeczne do kierunku jazdy) *Y* w strefie kontaktu koło/szyna. Na podstawie wyznaczonych sił w strefie kontaktu określony zostanie współczynniki wykolejenia *Y*/*Q*, jako maksymalna wartość stosunku siły bocznej do siły pionowej. Jeśli wartość współczynnika *Y*/*Q* nie przekroczy wartości 1,2, kryterium zagrożenia wykolejeniem pojazdu jest spełnione. To kryterium oparte na równowadze sił w nachylonej płaszczyźnie styku koła z szyną, a graniczna jego wartość przy zadanym profilu koła oraz przyjętym współczynniku tarcia wyznaczona jest z zależności (15).



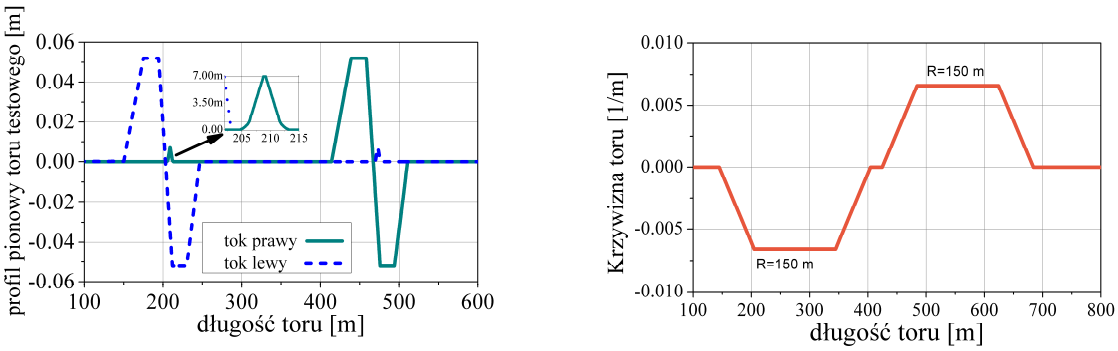
Rys. 12. Parametry geometrii kontaktu koła o profilu UIC 60 z szyną 49E1 w funkcji przemieszczenia poprzecznego y zestawu



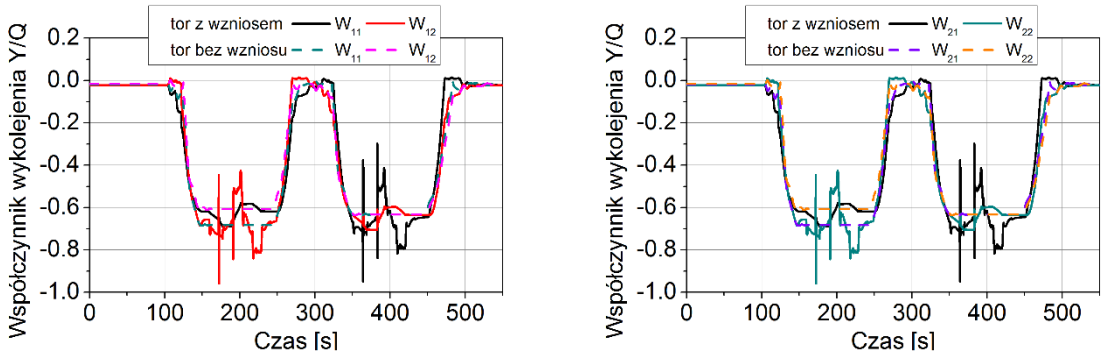
Rys. 13. Punkty kontaktu koła o profilu UIC 60 z szyną 49E1 podczas przemieszczenie poprzecznego tego koła

$$\frac{|Y|}{|Q|} < \frac{tg70 - 0{,}36}{1 + 0{,}36 \cdot tg70} \quad . \tag{15}$$

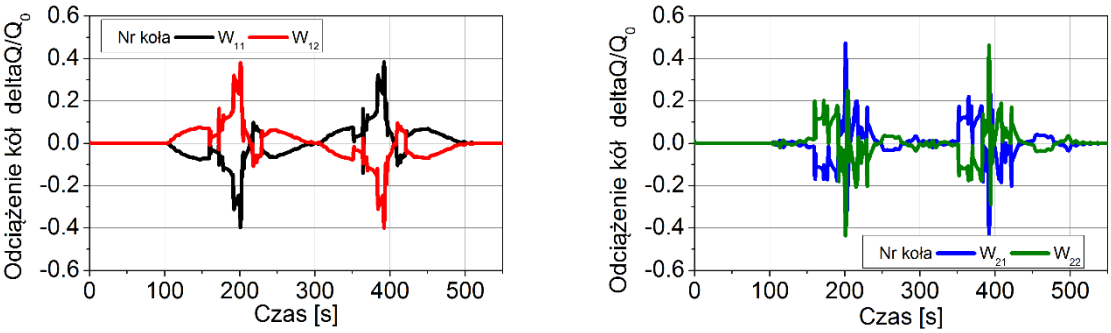
W przypadku, gdy otrzymana podczas badań dynamicznych wartość Y/Q>1,2, wówczas bezpieczeństwo przed wykolejeniem należy dodatkowo sprawdzić przez wznios koła. Wznios ten od pozycji zerowej nie powinien przekraczać Δz=5 mm. Podczas symulacji numerycznej założono prędkość przejazdu pojazdu równą v=5 km/h, zgodnie z badaniami eksperymentalnymi (metoda 3).



Rys. 14. Geometria zmodyfikowanego toru testowego z dodatkowym wichrowatością dla testu wózków



Rys. 15. Przebiegi współczynnika *Y/Q* kół podczas przejazdu po zmodyfikowanym torze testowym, zestaw prowadzący wózka pierwszego W_{11} i W_{12} i drugiego W_{21} i W_{22}



Rys. 16. Odciążenie kół zestawów prowadzących wózek pierwszy W_{11} i W_{12} i drugi W_{21} i W_{22} podczas przejazdu po zmodyfikowanym torze testowym

Pokazane na wykresach (Rys. 15) przebiegi wskaźnika wykolejenia ilustrują porównanie wyników przejazdu po torze bez dodatkowego wniosku (krzywe przerywane) i ze wzniosem odpowiadającym wichrowatości toru na bazie rozstawu zestawów kołowych pojedynczego wózka. Można zaobserwować, że wartości wskaźnika *Y/Q* w przypadku toru bez wzniosu mają wartości zbieżne z wynikami uzyskanymi eksperymentalne, tablica 2. W celu weryfikacji poprawności działania modelu, wyniki maksymalnych wartości współczynnika wykolejenia *Y/Q* otrzymane z podczas przejazdu po torze bez dodatkowego wzniosu zestawiono w tablicy 4. Pozwoliło to porównać testy eksperymentalne z numerycznymi, wynikiem czego jest wyznaczony moduł względnego błędu procentowego między nimi. Wartość tego błędu dla poszczególnych kół zawierała się w przedziale 3.03÷12,76%. W przypadku kół będących po stronie zewnętrznego toku szynowego w łuku toru błąd między wynikami jest większy, nie przekraczając jednak 13%. Uzyskany wynik świadczy o poprawnie

sformułowanym opisie cech modelu wagonu towarowego i jest podstawą do stwierdzenia prawidłowej walidacji modelu. Znacznie lepszą zgodność ilościową uzyskano przy wyznaczaniu momentu oporowego na czopie skrętu przez porównanie wyników z badań numerycznych i otrzymanych z testów eksperymentalnych opisanych w rozdziale 3 (Rys. 4). Stworzony model symulacyjny układu wagon –tor pozwala także na wyznaczenie odciążenia kół badanego pojazdu, co przedstawiono na rysunkach 16, które to stanowią także kluczowy wskaźnik weryfikacji bezpieczeństwa pojazdu przed wystąpieniem wykolejenia.

Tablica 4. Maksymalne wartości współczynnika wykolejenia z badań numerycznych oraz moduł błędu procentowego względem wyników z testów doświadczalnych

Koło W_{ij}	**11**	**12**	**21**	**22**	**31**	**32**	**41**	**42**
Y/Q [kN]	0.67	0.63	0.23	0.27	0.56	0.61	0.68	0.66
Błąd wzgl.| | [%]	7.89	12.76	9.52	12.50	5.48	11.46	3.03	6.45

5. Podsumowanie i wnioski

W artykule przybliżono zjawisko wykolejenia pojazdów szynowych. Sklasyfikowano metody oceny ryzyka wystąpienia tego zjawiska na podstawie przeglądu literatury i norm. Przedstawiony opis badań eksperymentalnych bezpieczeństwa przed wykolejeniem pokazał stosowane metodologie przy ich wykonywaniu. Uzyskane i opracowane wyniki z badań eksperymentalnych, przeprowadzonych wg wyżej opisanych metod badawczych, pozwoliły uwiarygodnić wyniki uzyskane z badań numeryczny powstałego modelu. Testy określające przebiegi wartości współczynników wykolejenia, odciążenia kół, momentów oporowych oparto na modelu numerycznym opisującym badany eksperymentalnie wagon towarowy. W pracy zaproponowano także nową, innowacyjną geometrię toru testowego do badania bezpieczeństwa przed wykolejeniem. Innowacja ta polegała na wprowadzeniu dodatkowego pionowego wzniosu w tokach szynowych toru stanowiąc wichrowatości (Rys. 14), która bazuje na rozstawie zestawów kołowych w pojedynczym wózku. Zaletą zastosowania w rzeczywistości takiej geometrii toru podczas badań rzeczywistych, także numerycznych, może pozwolić na ograniczenie kosztów badań i testów dopuszczeniowych pojazdu przez skrócenie czasu takich badań. Oszczędność ta wynika z faktu wykonania trzech przejazdów na tak zmodyfikowanym torze rzeczywistym, bez potrzeby dodatkowych testów wg innych metod oceny ryzyka i bezpieczeństwa przed wykolejeniem opisanych w 2 rozdziale tej pracy. Otrzymane przy użyciu nieliniowego modelu pojazdu kolejowego wyniki numeryczne pokazują, w jaki sposób na bezpieczeństwo jazdy mają wpływ różne czynniki związane z budową pojazdu i toru. Na podstawie przeprowadzonych badań wskazano, że wskaźnika *Y/Q* jest silnie uzależniony od wichrowatości wynikającej z bazy wózka. Analiza wyników uzyskanych z testów numerycznych nie wskazała przypadku wystąpienia ryzyka wykolejenia nawet przy zmodyfikowanym torze, co także potwierdziły badania eksperymentalne przeprowadzone wg opisanych metod normatywnych. Oszacowany na poziomie poniżej 13% błąd między wynikami eksperymentalnymi i teoretycznymi wskaźnika wykolejenia wskazuje dużą zgodności modelu teoretycznego z rzeczywistym pojazdem szynowym. Należy także stwierdzić, że równica kątów nabiegania zestawów kołowych pojedynczego wózka oraz obrót wózka względem pionowej osi ma znaczący wpływ na powtarzalność uzyskiwanych wyników.

6. Bibliografia

1. Bogacz R., Czyczuła W., Konowrocki R. Influence of sleepers shape and configuration on track-train dynamics, Shock and Vibration 2014; Article ID 393867-1-7: 8 pages.

2. Bogacz R., Frischmuth K., On dynamic effects of wheel–rail interaction in the case of Polygonalisation, Mechanical Systems and Signal Processing 2016; 79: 166-173.
3. Bogacz R., Konowrocki R. On new effects of wheel-rail interaction, Archive of Applied Mechanics 2012; 82: 1313-1323.
4. Bogdevicius M, Zygiene R. Research of dynamic processes of the system “vehicle – track” using the new method of vehicle wheel with metal scale. Eksploatacja i niezawodnosc – Maintenance and Reliability 2018; 20 (4): 638–649.
5. Bosso N., Gugliotta A., Somà A. Simulation of a freight bogie with friction damper, Politecnico di Torino, 5th ADAMS/Rail users conference, Harlem, The Netherlands - May 2000.
6. Chudzikiewicz A., Bogacz R., Kostrzewski M., Konowrocki R., Condition monitoring of railway track systems by using acceleration signals on wheelset axle-boxes, Transport 2018; 33; 2: 555-566.
7. Chudzikiewicz A., Opala M. Application of computer simulation methods for running safety assessment of railway vehicles in example of freight cars, Applied Mechanics and Materials 2008; 9: 61-69.
8. Dailydka S., Lingaitis L.P., Myamlin S., Prichodko V. Mathematical model of spatial fluctuations of passenger wagon. Eksploatacja i niezawodnosc – Maintenance and Reliability 2008; 4 (40): 4–8.
9. Dyniewicz B., Bajer C.I., Matej J. Mass splitting of train wheels in the numerical analysis of high speed train-track interactions, Vehicle System Dynamics 2015; 53(1): 51-67.
10. Elkins J., Wu H. New criteria for flange climb derailment, Proceedings of the ASME/IEEE Joint Railroad Conference 2000; 1-7.
11. EN 14363, Railway applications - Testing for the acceptance of running characteristics of railway vehicles - Testing of running behaviour and stationary tests, European Committee For Standardization, 2016.
12. Federal Railroad Administration. Track Safety Standards, Part 213. Subpart G. September, 1998.
13. Garcia de Jalon J., Bayo E. Kinematic and Dynamic Simulation of Multibody Systems. Springer-Verlag, 1994.
14. Garg V.K., Dukkipati R.V. Dynamics of Railway Vehicle Systems. Academic Press 1984
15. Gaspar P., Szabo Z., Bokor J. Observer based estimation of the wheel-rail friction coefficien. IEEE Conference on Computer Aided Control System Design 2006: 1043-1048.
16. Ge X., Wang K., Guo L., Yang M., Lv K, Zhai W. Investigation on derailment of empty wagons of long freight train during dynamic braking. Shock and Vibration 2018; 18 Article ID 2862143, 18 pages.
17. GosNIIV-VNIIZhT: Norms for analysis and design of railway wagons MPS 1520 mm (not self-propelled). 1996.
18. He J., Ben-Gera T., Liu X. Risk analysis of freight-train derailment caused by track geometry defect. ASME. ASME/IEEE Joint Rail Conference, 2016 *Joint Rail Conference*: V001T06A007. doi:10.1115/JRC2016-5743.
19. Hertz H. Über die berührung fester elastischer Körper (On the contact of rigid elastic solids). J. Reine und Angewandte Mathematik 1882; 92: 156-171.

20. Iwnicki S. (ed.) Handbook of Railway Vehicle Dynamics, CRC Press Inc., 2006.
21. Flammini F. Railway safety, reliability, and security: Technologies and Systems Engineering, IGI Global, 2012.
22. Kalker J.J. Three-dimensional elastic bodies in rolling contact. Springer, 1990.
23. Kalker J.J. A fast algorithm for the simplified theory of rolling contact, Vehicle System Dynamics 2007;11 (1): 1-13.
24. Kardas-Cinal E. Spectral distribution of derailment coefficient in non-linear model of railway vehicle–track system with random track irregularities. ASME. *J. Comput. Nonlinear Dynam.* 2013;8(3):031014-031014-9. doi:10.1115/1.4023352.
25. Koci, H.H., Swenson. C. A. Locomotive wheel-loading a system approach. General motors electromotive division. LaGrange, IL, February, 1978.
26. Konowrocki R., Bajer C.I. Friction rolling with lateral slip in rail vehicles, Journal of Theoretical And Applied Mechanics 2009; 47(2): 275-293.
27. Krishna, V.V., Berg, M., Stichel, S. Tolerable longitudinal forces for freight trains in tight S-curves using three-dimensional multi-body simulations. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 2019.
28. Krzyżyński T. On continuous subsystem modelling in the dynamic interaction problem of a train-track-system, Vehicle System Dynamics 2007; 24:311-324.
29. Liu X. Saat M.R., Barkan Ch. Freight-train derailment rates for railroad safety and risk analysis. Accident Analysis & Prevention 2017; 98: 1-9.
30. Matej J. Controlled wheel flange climb derailment of the load-measuring wheel set on curved and straight track, Proceedings of the Institute of Vehicles 2015;1(101): 75-90.
31. Matej J. A new mathematical model of the behaviour of a four-axle freight wagon with UIC single-link suspension. *Proceedings of the Institution of Mechanical Engineers Part F-Journal of Rail and Rapid Transit* 2011; 225(6): 637-647.
32. Matsudaira T. Dynamics of high speed rolling stock, Japanese National Railways RTRI Quarterly Reports, Special Issue, 1963.
33. Matsumoto A., et al. Continuous observation of wheel/rail contact forces in curved track and theoretical considerations, Vehicle System Dynamics 2019; 50(1): 349-364.
34. Miwa M., Oyama T. Modeling an optimal track maintenance and management strategy in consideration of train derailment accident risk, Journal of Japan Society of Civil Engineers 2019; 75(1): 11-28.
35. Okamoto I., Uchida M. The coefficient of friction in railway vehicles and tracks. Journal of Japanese Society of Tribologists 2002; 47(4): 249-254.
36. Opala, M. Evaluation of bogie centre bowl friction models in the context of safety against derailment simulation predictions, Archive of Applied Mechanics 2018; 88(6): 943–953.
37. Piotrowski J., Październiak P. Influence of dither generated by rolling contact on friction damping in freight wagons, Vehicle System Dynamics 2010; 48: 195-209.
38. PKP Polskie Linie Kolejowe S.A.: Instrukcja Id-1 (D1) - Warunki techniczne utrzymania nawierzchni na liniach kolejowych. Warszawa 2005.
39. Rausand M. Risk Assessment: Theory, Methods, and Applications, Wiley, 2011.

40. Raport ORE/ERRI B55 Rp.8 – Prevention of derailment of goods wagon on distorted tracks, 1983.

41. Reliability, Safety, and Security of Railway Systems. Modelling, Analysis, Verification, and Certification, Fantechi A., Lecomte T. Romanovsky A. (Eds.) Second International Conference, RSSRail 2017, Pistoia, Italy, November 14-16, 2017, Proceedings Series Springer Volume 10598, 2017.

42. Riessbeger K. Zur entgleisungssicherheit der rollenden landstrasse. ZEVRail Glasers Annalen. No 2/3 1994.

43. Sato E. et al., Lateral force during curve negotiation of forced steering bogies, Quarterly Report of RTRI 2003; 44(1): 8-14.

44. Shabana A. Dynamics of multibody systems. Cambridge University Press, Third Editon, 2005.

45. Szolc, T. Simulation of dynamic interaction between the railway bogie and the track in the medium frequency range. Multibody System Dynamics 2001;6(2): 99-122.

46. UIC 518, 2009, Testing and approval of railway vehicles from the point of view of their dynamic behaviour - Safety - Track fatigue - Ride quality.

47. VII001, AAR Mechanical Division, Manual of Standards and Recommended Practices. Section C- Part II, Volume 1, Chapter XI. Section 11.5.2 Track-Worthiness Criteria, Adopted 1987, Revised 1993.

48. Weinstock H. Wheel climb derailment criteria for evaluation of rail vehicle safety, Paper no. 84-WA/RT-1, 1984 ASME Winter Annual Meeting, New Orleans, LA, November, 1984.

49. Wu H., Wilson, N. Railway vehicle derailment and prevention. In: Iwnicki, S. ed. Handbook of Railway Vehicle Dynamics, Chapter 8, Taylor & Francis, Boca Raton, FL, 2006.

50. Zhou H., Zhang J., Hecht, M. Three-dimensional derailment analysis of crashed freight trains. Vehicle System Dynamics 2014; 52(3): 341-361.

Dr hab. inż. Tomasz Knefel

Dr hab. inż. Jacek Nowakowski

Katedra Silników Spalinowych I Pojazdów
Akademia Techniczno-Humanistyczna
43-309 Bielsko-Biała, ul. Willowa 2
E-mail: knefel@ath.bielsko.pl, jnowakow@ath.bielsko.pl

Modelowa analiza parametrów procesu wtrysku w układzie zasilania typu Common Rail

Keywords: *układ wtryskowy Common Rail, obliczenia symulacyjne, proces wtrysku, parametry eksploatacyjne układu zasilania*

Abstract: W pracy przedstawiono uproszczony model zasobnikowego układu zasilania w paliwo silnika o zapłonie samoczynnym. W rozważaniach nie uwzględniono przewodów wysokiego ciśnienia, a do symulacji pracy wtryskiwaczy opracowano empiryczny podmodel. Przedstawiono podstawowe równania modelu. Zjawiska zostały opisane układem 17 równań różniczkowych zwyczajnych, pierwszego rzędu. W pracy również zawarto ocenę wpływu parametrów geometrycznych zasobnika na proces wtrysku. Ocenę przeprowadzono za pomocą programu obliczającego proces wtrysku, wykorzystującego model rozważanego układu wtryskowego. Zaproponowano sposób wstępnego doboru parametrów geometrycznych zasobnika.

1. Wstęp

Rozwój współczesnych szybkoobrotowych silników o zapłonie samoczynnym jest związany z rozwojem ich układów zasilania w paliwo. Obecnie w zasilaniu tego rodzaju silników dominuje zasobnikowy układ wtryskowy Common Rail. Przy doborze układu do silnika należy uwzględnić wiele czynników eksploatacyjnych i regulacyjnych. Wykorzystanie badań symulacyjnych do analizy tych parametrów znacznie ułatwia i przyspiesza prowadzenie prac rozwojowych.

W zasobnikowym układzie zasilania w paliwo, wytwarzanie wysokiego ciśnienia następuje w tłoczkowej pompie wysokiego ciśnienia, skąd przepływa przewodem wysokiego ciśnienia do zasobnika, po czym przez stosunkowo krótkie przewody wtryskowe zostaje podane do wtryskiwaczy.

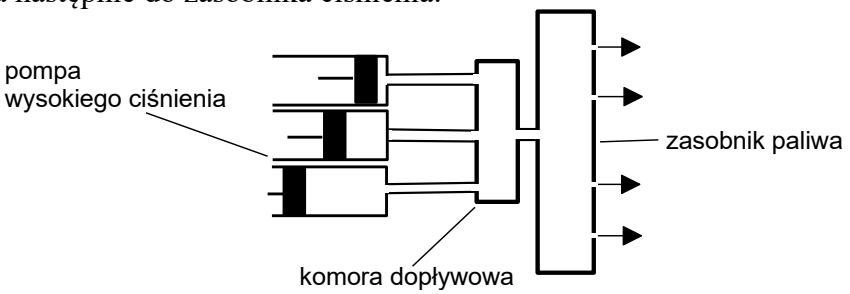
Istniejące modele układu wtryskowego typu Common Rail były tworzone przez zespoły badawcze przede wszystkim w celu porównania ich parametrów pracy i osiągów z parametrami innych systemów wtryskowych. Obejmowały one analizy i dyskusje podstawowych czynników związanych z przebiegiem wtrysku [3], lecz również pozwalały wyznaczać wielkości, które trudno zmierzyć, na przykład efektywne pole powierzchnię przepływu. Można również spotkać rozważania dotyczące elementów układów zasilania silników o zapłonie samoczynnym i zagadnień związanych z ich sterowaniem. Jedną z podstawowych jest tutaj praca [2], w której autorzy jako jedni z pierwszych podjęli problem przepływu paliwa i sterowania wtryskiwaczem. Z kolei w pracy [4] opracowano model zaworu przelewowego sterującego ciśnieniem paliwa oraz model zaworu dławiącego dopływ paliwa do pompy wysokiego ciśnienia, które rozważano wraz podmodelem tej pompy. Ta

ostatnia była sterowana pseudolosową sekwencją bitów. Przeprowadzano również symulacyjną analizę właściwości materiałów używanych do wykonania elektrozaworów wtryskiwaczy i pracy zaworów [8]. Z kolei autorzy pracy [10] w swoich rozważaniach skupili się na sztywności zespołu tłok sterujący-iglica, wartościach współczynników wypływu otworków rozpylacza i wyznaczeniu bezwymiarowej liczby kawitacyjnej. Określenie wpływu zjawisk falowych w zasobniku na przebieg wtrysku było głównym tematem pracy [1] wykorzystanej w koncernie Daimler Chrysler AG oraz [6], gdzie oceniano wpływ własności i ciśnienia paliwa, czasu wtrysku oraz długości i średnicy przewodu wtryskowego na zmiany ciśnienia w systemie. Z kolei ocenę oraz dobór wymiarów geometrycznych systemu wtryskowego zawarto w pracach [1, 9]. Inną grupę stanowią prace, które opisują modele ukierunkowane na sterowanie ciśnieniem w zasobnikowym systemie wtryskowym. Tu brano pod uwagę moduł sprężystości, ciśnienie, temperaturę paliwa oraz prędkość obrotową silnika, a efektami były zlinearyzowane modele sterowania oraz wstępne konstrukcje sterownika czy regulatora, służących do atestacji systemu sterowania ciśnienia w zasobniku [5, 7]. Dalej wykracza praca [11] przedstawiająca zasobnikowy system wtryskowy czwartej generacji i model on-line korygujący przebiegi wypływu paliwa z rozpylacza. Stanowiła inspirację do opracowania uproszczonego modelu układu zasilania.

Daje się zauważyć tendencję do stosowania oprogramowania do modelowania i analizy układów jednowymiarowych, wielodziedzinowych, mechatronicznych (interface, analizy statyczne i dynamiczne). W większości przypadków stosowane są pakiety AMESim oraz Matlab/Simulink. Nie jest to jednak regułą. Często wstępne opracowania są wykonywane za pomocą tych pakietów, a dokładniejsze, dedykowane określonemu celowi, z wykorzystaniem konwencjonalnych języków programowania. Zazwyczaj są to właśnie jednowymiarowe modele, opisujące nieustalony, sprężysty przepływ paliwa w układzie. Jednak badania symulacyjne prowadzi się w oparciu o modele obliczeniowe o różnym stopniu złożoności. Często w analizach technicznych są stosowane uproszczone modele, z uwagi na mniejszą liczbę czynników wpływających na przebieg analizowanych procesów. W takich przypadkach należy zawsze określić wpływ uproszczeń na dokładność odwzorowania rozważanych zjawisk.

Po przeprowadzeniu analizy dostępnej literatury stwierdzono, że brak modeli wykorzystujących równania mechaniki płynów w powiązaniu z wynikami badań stanowiskowych. Dlatego zdecydowano o podjęciu prac nad teoretyczno-empirycznym modelem układu, z wykorzystaniem konwencjonalnego języka programowania.

Przedstawiona praca dotyczy uproszczonego modelu układu, w którym paliwo z pompy wysokiego ciśnienia o trzech tłokach, jest podawane do zbiorczej komory dopływowej, a następnie do zasobnika ciśnienia.



Rys. 1. Schemat modelowanego układu zasilania

W rozważaniach nie uwzględniono przewodu wysokiego ciśnienia, przewodów wtryskowych i wtryskiwaczy. Wtrysk (wypływ paliwa) następuje przez cztery otwory, bezpośrednio z zasobnika ciśnienia (rys. 1).

2. Równania różniczkowe opisujące układ

Równania różniczkowe rozważanego układu można zapisać w postaci przedstawionej poniżej.

2.1. Równania określające ciśnienie w komorach pompy

Z uwagi na niewielkie wymiary komór przyjęto, że zmiany ciśnienia $p_p^{(i)}$ w czasie można wyznaczyć z uproszczonych równań ciągłości:

$$\frac{dp_p^{(i)}}{dt} = \frac{E_p^{(i)}}{V_p^{(i)}}\Big[\, A_p^{(i)} \frac{dh_p^{(i)}}{dt} - sgn(p_p^{(i)}\text{-}p_d)\,\varepsilon_A^{(i)}\,\mu_d F_d^{(i)} \sqrt{\frac{2}{\rho_p^{(i)}}\left|p_p^{(i)} - p_d\right|} + \\ - sgn\,(p_p^{(i)}\text{-}p_L)\,\varepsilon_B^{(i)}\,\mu_w F_w^{(i)} \sqrt{\frac{2}{\rho_p^{(i)}}\left|p_p^{(i)} - p_L\right|} + -\varepsilon_u^{(i)} F_u^{(i)} \sqrt{\frac{2}{\rho_z^{(i)}}\left|p_p^{(i)} - p_z\right|}\,\Big]\eta_p\,\eta_t \tag{1}$$

dla i = 1, 2, 3.

gdzie: $p_p^{(i)}$ – ciśnienie w komorze *i*-tego tłoka pompy,

$V_p^{(i)}$ –objętość komory *i*-tego tłoka pompy,

$E_p^{(i)} = E_p^{(i)}(p_p^{(i)}, T)$ - moduł sprężystości paliwa w komorze *i*-tego tłoka,

$A_p^{(i)}$ – pole powierzchni *i*-tego tłoka pompy,

$\frac{dh_p^{(i)}}{dt}$ - prędkość *i*-tego tłoka pompy,

p_d – ciśnienie zasilania,

μ_d – współczynnik przepływu otworu dopływowego,

$\varepsilon_A^{(i)}$ – wskaźnik sterujący,

$F_d^{(i)}$ – pole powierzchni otworu dopływowego do komory *i*-tego tłoka pompy,

$\rho_p^{(i)} = \rho(\rho_p^{(i)}, T)$ – gęstość paliwa w komorze *i*-tego tłoka,

$\varepsilon_B^{(i)}$ – wskaźnik sterujący,

p_L – ciśnienie w komorze dopływowej,

μ_w – współczynnik przepływu otworu wypływowego,

$F_w^{(i)}$ – pole powierzchni otworu wypływowego z komory i-tego tłoka pompy,

$\varepsilon_u^{(i)}$ – wskaźnik sterujący,

$F_u^{(i)}$ – pole powierzchni otworu upustowego,

p_z – ciśnienie w zasobniku,

η_p – sprawność pompy zależna od prędkości obrotowej i ciśnienia paliwa,

η_t – poprawka uwzględniająca zmianę sprawności *i*-tego tłoka pompy wysokiego ciśnienia w zależności od temperatury paliwa.

W powyższym wzorze wprowadzono współczynniki sterujące, których interpretacja jest następująca:

$\varepsilon_A^{(i)}$ - człon wydatku objętościowego, zależnego od różnicy ciśnień $p_p^{(i)}$ i p_d jest aktywny tylko wówczas, gdy wznios grzybka zaworu dopływowego $h_g^{(i)} \rangle 0$,

$\varepsilon_B^{(i)}$ - człon wydatku objętościowego, zależnego od różnicy ciśnień $p_p^{(i)}$ i p_L jest aktywny tylko wówczas, gdy wznios kulki zaworu łączącego komorę pompy z komorą dopływową $h_k^{(i)}$ jest większy od zera;

$\varepsilon_u^{(i)}$ – trzeci człon wypływowy jest uaktywniany w równaniu (1) wówczas, gdy ciśnienie w zasobniku p_z przekracza przyjęte ciśnienie graniczne $p_z^{(gr)}$, równocześnie tłok pompy porusza się do góry ($\dot{h}_p^{(i)} \rangle\, 0$) oraz błąd względny odchylenia p_z od $p_z^{(gr)}$ przekracza dopuszczalną wartość ε.

Ponadto przyjmowano, że przekrój upustu $F_u^{(i)}$ zmienia się, zależnie od wartości różnicy ciśnień p_z i $p_z^{(gr)}$, według wzoru:

$$F_u^{(i)} = F_{u0}^{(i)} \sqrt{\frac{2}{\rho_z}\left|p_z - p_z^{gr}\right|} \tag{2}$$

Równania (1) są równaniami różniczkowymi zwyczajnymi, pierwszego rzędu, nieliniowymi.

2.2. Równanie ruchu grzybkowych zaworów dopływowych

Z drugiej zasady dynamiki Newtona wynika, że prostoliniowy ruch zaworów dopływowych opisują równania:

$$\frac{d^2 h_g^{(i)}}{dt^2} = \frac{\varepsilon_g^{(i)}}{m_g^{(i)}} f_g^{(i)} \tag{3}$$

gdzie: $f_g^{(i)} = -\left(h_g^{(i)} + h_{gl0}^{(i)}\right)k_g^{(i)} - \left(p_p^{(i)} + p_d^{(i)}\right)F_g^{(i)}$

$F_g^{(i)} = \dfrac{\pi\left[g_1^{(i)}\right]^2}{4}$ dla i = 1, 2, 3,

$h_g^{(i)}$ – wznios grzybka zaworu dopływowego,

$m_g^{(i)}$ – masa grzybka,

$h_{gl0}^{(i)}$ – wstępne napięcie sprężyny,

$k_g^{(i)}$ – stała sprężyny,

$F_g^{(i)}$ – pole powierzchni zaworu.

Występujący w powyższych równaniach wskaźnik sterujący εg przyjmuje wartości:

$$\varepsilon_g^{(i)} = \begin{cases} 1 & \text{gdy } wG^{(i)} = 1 \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

Z kolei wskaźnik $wG^{(i)}$ wskazuje, czy grzybek zaworu dolotowego osiadł na gnieździe ($wG^{(i)}$ = 0), jest w fazie ruchu ($wG^{(i)}$ = 1), czy też osiągnął maksymalny wznios $h_{g_{\max}}^{(i)}$ ($wG^{(i)}$ = 2), przy czym $wG^{(i)} \in \{0, 1, 2\}$.

Równania (3) stanowią układ równań trzech równań różniczkowych zwyczajnych, drugiego rzędu.

2.3. Równanie ruchu kulowych zaworów wypływowych

Równania ruchu zaworów wypływowych mają postaci:

$$\frac{d^2 h_k^{(i)}}{dt^2} = \frac{\varepsilon_k^{(i)}}{m_k^{(i)}} f_k^{(i)} \tag{4}$$

gdzie: $f_k^{(i)} = -\left(h_k^{(i)} + h_{kl0}^{(i)}\right)k_k^{(i)} - \left(p_p^{(i)} + p_o^{(i)}\right)F_w^{(i)}$

dla i = 1, 2, 3,

$h_k^{(i)}$ – wznios kuli zaworu wypływowego,

$m_k^{(i)}$ – masa kuli,

$h_{kl0}^{(i)}$ – wstępne napięcie sprężyny,

$k_k^{(i)}$ – stała sprężyny,

$F_w^{(i)}$ – pole powierzchni otworu wypływowego.

Występujący równaniach wskaźnik sterujący εk przyjmuje wartości:

$$\varepsilon_k^{(i)} = \begin{cases} 1 & \text{gdy} \quad wK^{(i)} = 1 \\ 0 & \text{w przeciwnym przypadku} \end{cases} \qquad wK^{(i)} \in \{0, \quad 1, \quad 2\}$$

Z kolei wskaźnik $wK^{(i)}$, podobnie jak wskaźnik $wG^{(i)}$, wskazuje pozycje kuli zaworu: 0 jeśli brak przepływu między komorą pompy i komorą dopływową, 1 jeśli kulka jest w fazie ruchu, 2 jeśli kulka osiągnęła maksymalny wznios.

Równania (4) stanowią układ równań trzech równań różniczkowych zwyczajnych, drugiego rzędu.

2.4. Równanie określające ciśnienie w komorze dopływowej

Podobnie jak w przypadku komór pompy wysokiego ciśnienia przyjęto, że zmiany ciśnienia w komorze dopływowej można wyznaczyć z uproszczonego równania ciągłości:

$$\frac{dp_L^{(i)}}{dt} = \frac{E_L^{(i)}}{V_L^{(i)}} \Big[\operatorname{sgn}(p_p^{(i)} - p_L)\, \varepsilon_B^{(i)}\, \mu_w^{(i)} F_w^{(i)} \sqrt{\frac{2}{\rho_p^{(i)}} \left| p_p^{(i)} - p_L \right|} + \\ - \operatorname{sgn}(p_L - p_Z)\, \mu_L F_L \sqrt{\frac{2}{\rho_0} \left| p_L - p_Z \right|} \Big] \tag{5}$$

dla i = 1, 2, 3,

gdzie: $V_L^{(i)}$ – objętość komory dopływowej,

$\mu_w^{(i)}$ – współczynnik przepływu otworu dopływowego do zasobnika,

F_L – pole powierzchni otworu dopływowego do zasobnika, równe polu powierzchni przekroju przewodu łączącego komorę dopływową z zasobnikiem,

pozostałe oznaczenia jak w p. 2.1.

Warto zaznaczyć, że objętość V_L należy powiększyć o objętość przewodu, łączącego komorę dopływową z zasobnikiem:

$$V_L := V_L + \frac{\pi d^2}{4} L \tag{6}$$

gdzie: d – średnica przewodu, L – długość przewodu łączącego komorę dopływową z zasobnikiem.

2.5. Równania ciśnienia w zasobniku

Do opisu zmian ciśnienia w zasobniku również wykorzystano równanie ciągłości:

$$\frac{dp_z}{dt} = \frac{E_z}{V_z} \Big[-\sum_{i=1}^{4} \operatorname{sgn}(p_z - p_k)\, \varepsilon_z^{(i)}\, \mu_z^{(i)} A_z^{(i)} \sqrt{\frac{2}{\rho_z} \left| p_z - p_k \right|} + \\ \operatorname{sgn}(p_L - p_Z)\, \mu_L F_L \sqrt{\frac{2}{\rho_L} \left| p_L - p_Z \right|} \Big] \tag{7}$$

gdzie: $A_z^{(i)}$ – zmienne pole powierzchni wypływu paliwa przez otwór wypływowy,

V_L – objętość komory dopływowej,

p_k – ciśnienie w komorze spalania (przeciwciśnienie).

$$\varepsilon_z^{(i)} = \begin{cases} 1 & \text{gdy} \quad t \in \left\langle t_A^{(i)}, \, t_B^{(i)} \right\rangle \quad i \quad p_z > p_k \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

$t_A^{(i)}, t_B^{(i)}$ – czasy otwarcia otworów,

pozostałe oznaczenia jak w p. 2.1 i 2.4.

Równanie (7) jest równaniem różniczkowym zwyczajnym pierwszego rzędu, nieliniowym.

Modelowanie zjawisk hydrodynamicznych we wtryskiwaczu napotyka na szereg trudności. Zasadniczą sprawą dla właściwego modelu zasobnika jest określenie reguł, według których paliwo wypływa z zasobnika do komory spalania. W prezentowanym algorytmie

przyjęto, że kolejne otwory otwierają się co 180° obrotu wału pompy. Istotny jest również dobór wartości szeregu współczynników koniecznych do przeprowadzenia właściwej ilościowej oceny zachodzących zjawisk. Można tu wymienić współczynniki oporów hydraulicznych, współczynniki natężeń przepływu czy współczynniki oporów ruchu elementów ruchomych. Wartości tych wielkości są zmienne w zależności od ciśnienia paliwa, co utrudnia ich wyznaczanie. Ponadto przy modelowaniu elektronicznie sterowanych układów wtryskowych należy uwzględnić elektrozawory sterujące, które wymagają znajomości kolejnych wielkości, zwłaszcza własności materiałowych. Wartości niektórych wielkości są niekiedy trudne do oszacowania, dlatego zdecydowano o opracowaniu empirycznego modelu wypływu paliwa z rozpylacza, opartego o funkcję $A_z^{(i)}$. Wykorzystano wartości charakterystycznych czasów, uzyskane z analiz przebiegów wtrysku.

Podstawową obserwacją poczynioną podczas eksperymentów było stwierdzenie, że rzeczywisty przebieg wzniosu iglicy, a tym samym funkcji $A_z^{(i)}$ odbiega od teoretycznego, w którym określa się:

$t_B^{(i)}$ – zadany czas otwarcia, $t_0^{(i)}$ – zadany czas przerwy.

Przede wszystkim stwierdzono, że rzeczywisty czas otwarcia $\bar{t}_B$ jest większy od zadanego $t_B^{(i)}$ o wielkość w przybliżeniu stałą, oznaczoną jako $t_d^{(i)}$ – czas opóźnienia wtrysku. Czas opóźnienia uwzględnia różnice między zadanym, a realizowanym czasem wtrysku. Został wyznaczony doświadczalnie. Również przebiegi funkcji $A_z^{(i)}$ miały kształt bardziej zbliżony do paraboli niż do przebiegu teoretycznego, w postaci funkcji prostokątnej.

W zależności od wartości $t_0^{(i)}$ oraz $t_d^{(i)}$ otrzymuje się dwa różne przypadki:

$t_d^{(i)} < t_0^{(i)}$ oraz $t_d^{(i)} > t_0^{(i)}$

uwzględnione w opracowanym programie komputerowym.

W programie obliczającym proces wtrysku wprowadzono możliwość zadawania ciśnienia, poniżej którego wtrysk nie powinien się rozpocząć (jest to odpowiednik ciśnienia otwarcia wtryskiwaczy). Jest to zabezpieczenie przed obliczaniem parametrów wtrysku w przypadku, gdy jakość procesu rozpylenia (nie analizowana za pomocą tego modelu) mogłaby okazać się niezadowalająca.

3. Numeryczne całkowanie równań różniczkowych układu

Większość metod całkowania układów równań różniczkowych zwyczajnych wymaga sprowadzenia równań wyższego rzędu do równań pierwszego rzędu. Dlatego równania (3) i (4) sprowadzono do odpowiednich dwóch równań pierwszego rzędu. Równania (1), (3), (4) i (7) zapisano zatem w postaci układu równań 1-go rzędu postaci:

$$\dot{X} = F(t, X) \tag{8}$$

gdzie F jest funkcją wektorową a X:

$$X = \left[p_L, p_p^{(1)}, ..., p_p^{(3)}, h_g^{(1)}, \dot{h}_g^{(1)}, ..., h_g^{(3)}, \dot{h}_g^{(3)}, h_k^{(1)}, \dot{h}_k^{(1)}, ..., h_k^{(3)}, \dot{h}_k^{(3)}, p_z\right]^T$$

jest wektorem o $m = 17$ składowych.

Należy zatem całkować układ m = 17 równań różniczkowych zwyczajnych, pierwszego rzędu. Zastosowano do tego celu metodę Rungego-Kutty IV rzędu ze stałym krokiem całkowania.

4. Warunki początkowe, uwagi dodatkowe

Obliczenia prowadzono przyjmując, że w chwili początkowej (t = 0) wszystkie ciśnienia są równe ciśnieniu paliwa dopływającego (zasilania) p_d oraz zerowe są wzniosy i prędkości, to znaczy:

$$X_{t=0} = \left[p_d, p_d, p_d, p_d, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, p_d\right]^T \tag{9}$$

Przyjmowano też, że pracę rozpoczyna pierwsza sekcja pompy wysokiego ciśnienia, a pozostałe są uruchamiane po odpowiednio: 120 i 240 stopniach obrotu wału. Przyjmowano zatem, że kąty $\alpha^{(i)}$ są określone zależnościami:

$$\alpha^{(i)} = \begin{cases} 0 & \text{gdy } t < t_i \\ \omega(t\text{-}t_i) & \text{gdy } t > t_i \end{cases} \tag{10}$$

gdzie: ω - prędkość kątowa,

$$t_i = 0 \quad \text{dla } i = 1, \qquad t_i = \frac{2\pi}{3\omega} \quad \text{dla } i = 2, \qquad t_i = \frac{4\pi}{3\omega} \quad \text{dla } i = 3.$$

Ponieważ badania empiryczne prowadzono dla ustalonych warunków pracy analizowanego układu zasilania, wyniki otrzymane w wyniku symulacji komputerowych można uważać za właściwe dopiero po kilku cyklach pracy (φ > 720°), bowiem w początkowej fazie obliczeń zbyt wyraźny jest wpływ warunków początkowych (9) oraz przesunięć t_i we wzorze (10).

Jak wspomniano już poprzednio, otwory wypływowe w zasobniku uruchamiano kolejno (cyklicznie) co:

$$\Delta T = t_B^{(i)} - t_A^{(i)} = \frac{\pi}{\omega} \tag{11}$$

Program do modelowania pracy systemu wtrysku typu Common Rail opracowany w oparciu o wyżej przedstawione zależności umożliwia obliczenie przebiegów ciśnienia w komorach pompy, komorze dopływowej i zasobniku ciśnienia, wzniosów tłoków i ruchomych elementów zaworów. Obliczane są sumaryczne dawki wtrysku oraz natężenia wypływu paliwa przez poszczególne otwory wtryskowe. Przebiegi wtrysku można wyznaczyć dla niedzielonej i dzielonej dawki oraz różnych wartości czasu przerwy.

Obliczenia weryfikacyjne przeprowadzono dla układu zasilania z walcowym zasobnikiem wysokiego ciśnienia. Porównania dokonano dla: dzielonej dawki wtrysku - część pilotująca 450 μs, przerwa 600 μs, główna część 450 μs, zadanego ciśnienia w zasobniku na poziomie 700 barów, prędkości obrotowej pompy 695 obr/min, kolejności wtryskiwania 1 – 2 – 3 – 4. Różnice wartości obliczonych i zmierzonych dawek wtrysku wynosiły od 2,4 do 7,7 %, w zależności od grupy selekcyjnej wtryskiwacza. Wynikają one głównie z przyjętych uproszczeń w modelu, bowiem nie uwzględniono przewodów wysokiego ciśnienia i zespołów wtryskiwaczy. Znaczący wpływ na dawkę ma czas opóźnienia wtrysku i jego zależność od ciśnienia paliwa.

5. Wpływ parametrów geometrycznych zasobnika na proces wtrysku

Wykorzystując model rozważanego układu wtryskowego, za pomocą programu obliczającego proces wtrysku, wykonano obliczenia dla różnych wartości zadanego sygnału sterującego wtryskiwaczem. Rozważano podawanie dwuczęściowej dawki paliwa. Obliczenia miały na celu jakościową i ilościową ocenę wpływu badanych wielkości na parametry wtrysku. Przedstawiono rozważania dotyczące wpływu parametrów geometrycznych zasobnika paliwa na proces wtrysku.

Zasobnik jest konstrukcyjnie stosunkowo prostym elementem, jednak pełni istotną rolę w ograniczeniu propagacji fal ciśnienia. Odpowiednio dobrana objętość zapewnia

ciągłość dawkowania przy gwałtownych zmianach parametrów pracy silnika. Jak wcześniej wspomniano, do przeprowadzania modelowych obliczeń przyjęto walcowy zasobnik wysokiego ciśnienia układu wtryskowego silnika o zapłonie samoczynnym klasy 1700 cm^3.

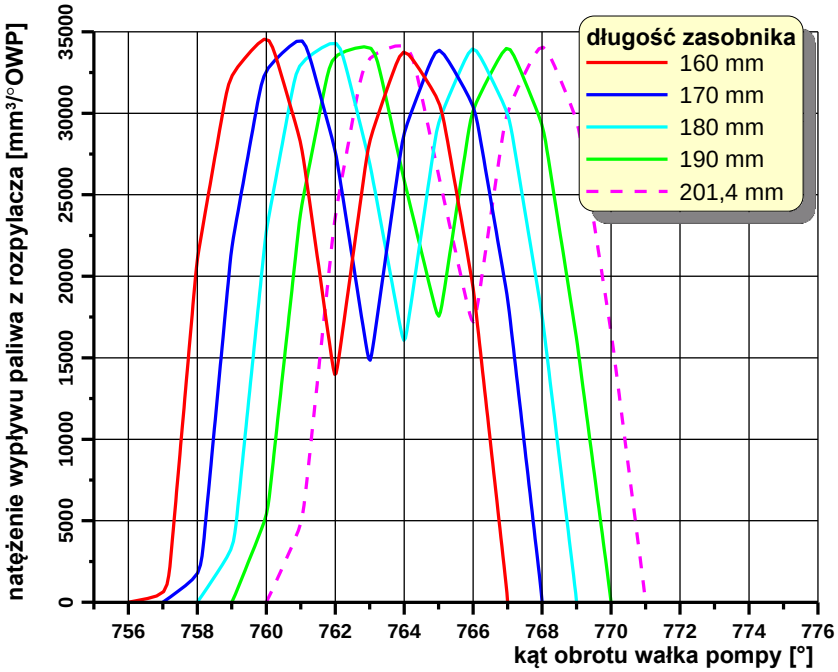
Za pomocą modelu oceniono wpływ długości, średnicy i objętości zasobnika na parametry procesu podawania paliwa. Obliczenia wykonano: dla stałej średnicy zasobnika i zmiennej długości oraz stałej długości i zmiennej średnicy zasobnika. Analizowano zmiany przebiegu procesu wtrysku, dawki paliwa, kąta początku wtrysku oraz kąta trwania wtrysku.

Ocena wpływu długości zasobnika

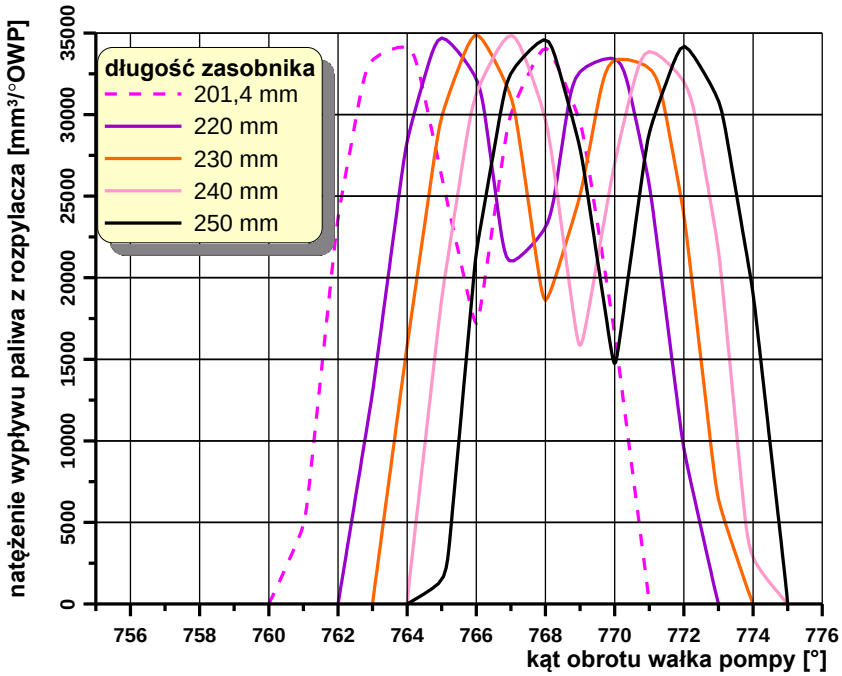
Na rys. 2 i 3 przestawiono przebiegi wtrysku jednego z wtryskiwaczy obliczone dla stałej średnicy i różnych długości zasobnika. Fioletową przerywaną linią przedstawiono wyniki dla podstawowej, zastosowanej przez producenta, długości zasobnika 201,4 mm. Tutaj kąt wtrysku wynosi 11° i pozostaje taki sam dla wszystkich przypadków. Zmienia się natomiast kąt początku wtrysku i dla rozważanego zakresu długości zasobnika zakres zmian wynosi 8° OWP. Jest dość istotna zmiana ważnego parametru wtrysku, którą należy uwzględniać przy projektowaniu algorytmów sterujących pracą silnika. Zmiany te przede wszystkim wynikają ze sposobu sterowania wtryskiwaczem w modelu, który umożliwia jego otwarcie przy zadanej wartości ciśnienia.

Wraz ze zwiększaniem długości zasobnika, średnie wartości ciśnienia w zasobniku zmieniają się zaledwie o 0,02 % i te zmiany są praktycznie niezauważalne. Podobnie nieznacznym zmianom ulegają ekstrema natężenia wypływu paliwa z rozpylacza.

Natomiast zmieniają się różnice między maksymalną i minimalną wartością ciśnienia. Jeśli dla zasobnika o długości 160 mm różnica wynosi 77 barów, a dla 201,4 mm 62,6 bara, to dla 250 mm już tylko 51,9 bara. Te zmiany wpływają na zachowanie się paliwa w zasobniku.

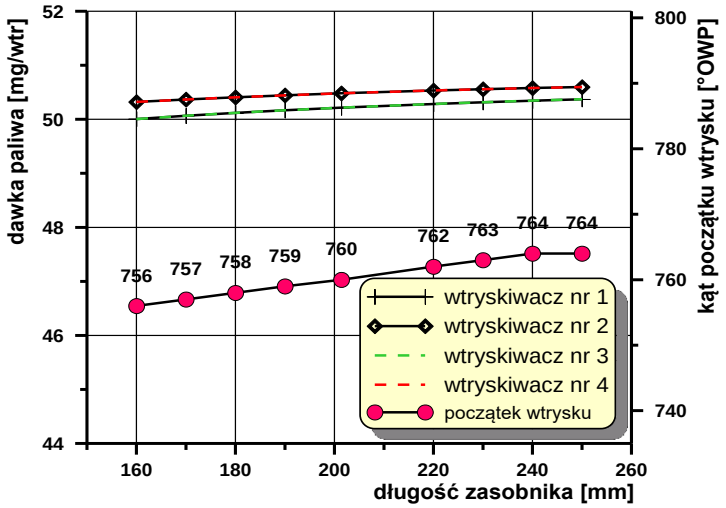


Rys. 2. Obliczone przebiegi wtrysku dla długości zasobnika 160 mm ÷ 201,4 mm

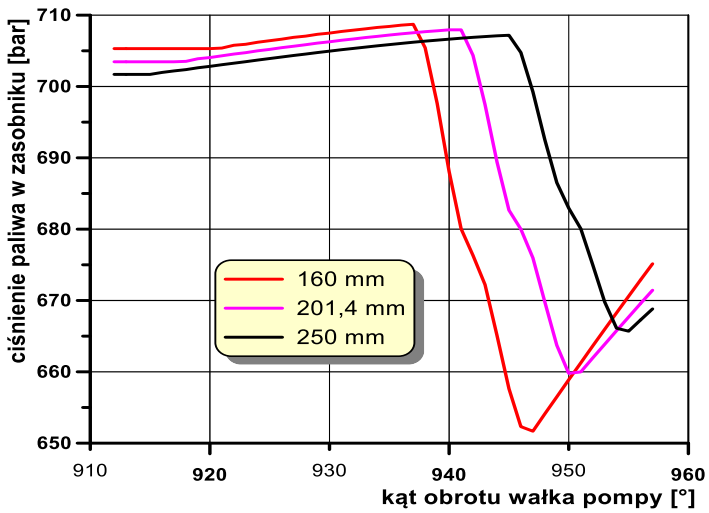


Rys. 3. Obliczone przebiegi wtrysku dla długości zasobnika 201,4 mm ÷ 250 mm

Przedstawione zmiany parametrów procesu wtrysku wynikają przede wszystkim ze zwiększania objętości rozważanego elementu. Ponieważ obliczenia prowadzono przy niezmienionych nastawach sterujących układem, zwiększenie objętości skutkuje coraz późniejszym osiąganiem wymaganego poziomu ciśnienia. Stąd opóźnienia początku kąta wtrysku (rys. 4). Ponieważ czas trwania wtrysku nie zmienia się, również coraz później następuje koniec wtrysku, który przypada dla coraz mniejszych różnic ciśnień między początkiem, a końcem wtrysku (rys. 5). Dlatego odnotowano niewielkie, bo wynoszące 0,7 %, zwiększenie dawki wtrysku.



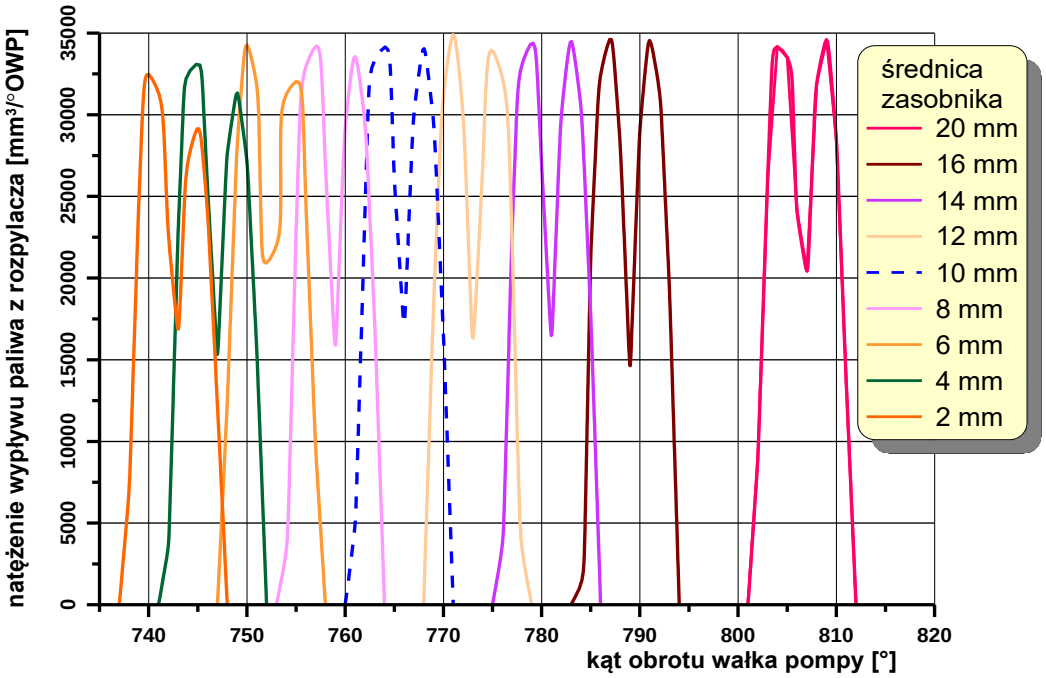
Rys. 4. Obliczone wartości dawek i kąta początku wtrysku dla różnych długości zasobnika



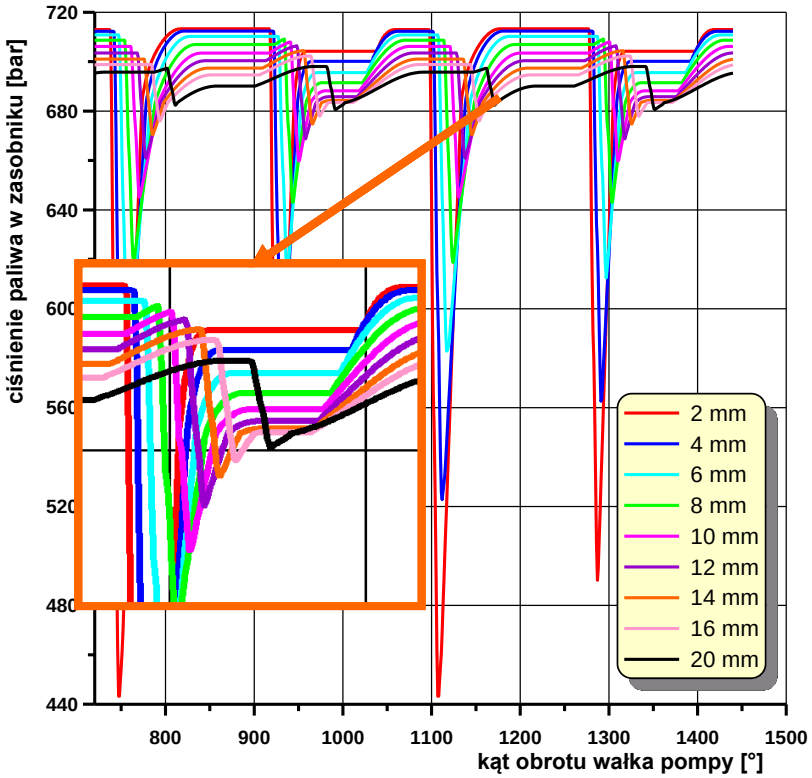
Rys. 5. Obliczone zmiany ciśnienia w zasobniku podczas wtrysku oraz kąty początku wtrysku dla wybranych długości zasobnika

Ocena wpływu średnicy zasobnika

Analizowane i zadawane wyżej zmiany długości zasobnika wysokiego ciśnienia miały liniowy charakter i tak też zmieniały się parametry wtrysku. Nieco inaczej jest, gdy bierze się pod uwagę przebiegi obliczone dla stałej długości i zmieniającej się średnicy zasobnika. W tych rozważaniach kierunek zmian jest podobny, jednak zmiany objętości są znaczne, a następują nieliniowo, zgodnie z drugą potęgą zadawanej średnicy. W celu pełniejszego zobrazowania zmian parametrów wtrysku przyjęto szeroki zakres zmian średnic, od najmniejszej odpowiadającej średnicy przewodu wtryskowego, do 20 mm, a więc wartości większej od stosowanych w większości zasobników samochodów osobowych. W odniesieniu do zmian długości zasobnika widać, że większa część energii dostarczonej do zasobnika jest zużywana w procesie ściskania cieczy. Zwiększona ilość paliwa w zasobniku przejmując część dostarczonej energii, powoduje znaczące opóźnienie początku wtrysku wynoszące aż 64° (rys. 6, a także rys. 8), przy niezmiennych wartościach kąta wtrysku wynoszących 11°. Jednak i w tym przypadku istotną rolę odgrywa sposób sterowania otwarciem wtryskiwacza. Średnie wartości ciśnienia w zasobniku zmieniają się o 1,5 % i nie oddają zmian następujących w zasobniku podczas procesu wtrysku paliwa, zwłaszcza dla najmniejszych średnic, gdzie mają miejsce duże zmiany ciśnienia (rys. 7). Stąd większe zmiany dawki i kąta początku wtrysku. Wraz ze zwiększaniem średnicy zasobnika jednak nie następuje znaczące zróżnicowanie natężeń wypływu paliwa z rozpylacza (rys. 6). Można je zauważyć jedynie dla najmniejszych średnic zasobnika, a więc tam, gdzie mają miejsce największe spadki ciśnienia. Forma wtrysku nie ulega zmianom. Różnice ciśnień (rys. 7) wpływają na zmianę ilości podawanego paliwa.

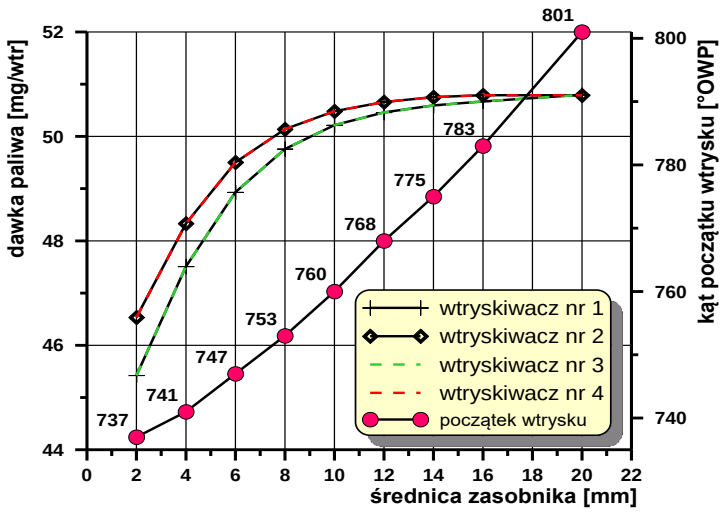


Rys. 6. Obliczone przebiegi wtrysku dla różnych średnic zasobnika



Rys. 7. Obliczone przebiegi ciśnienia dla różnych średnic zasobnika

Przy dużych spadkach ciśnienia, część procesu jest realizowana przy niskich wartościach ciśnienia, stąd mniejsza ilość paliwa (rys. 8). W całym rozważanym zakresie zmian średnic dawka wzrosła znacząco, bo o 11,8 %.



Rys. 8. Obliczone wartości dawek i kąta początku wtrysku dla różnych średnic zasobnika

Przedstawione wyżej wyniki modelowych analiz nie wyczerpują całości zagadnienia. Wykonano dodatkowe obliczenia, których wyniki uwzględniono w jakościowej ocenie wpływu omawianych parametrów pracy zasobnikowego układu wtryskowego na proces wtrysku (tab. 1).

Tabela 1. Ocena jakościowa badanych parametrów

Parametr	zakres wartości	wpływ badanego parametru na:		
		dawkę paliwa [mg]	kąt trwania wtrysku [°]	kąt początku wtrysku [°]
czas przerwy	200 µs÷900 µs	+	+ + +	- -
czas opóźnienia wtrysku	100 µs÷900 µs	+ + +	+ + +	- -
długość zasobnika (proporcjonalnie zmieniany rozstaw króćców)	160 mm÷250 mm	+ -	- -	+
średnica zasobnika	2 mm÷20 mm	+ + +	- -	+ + +

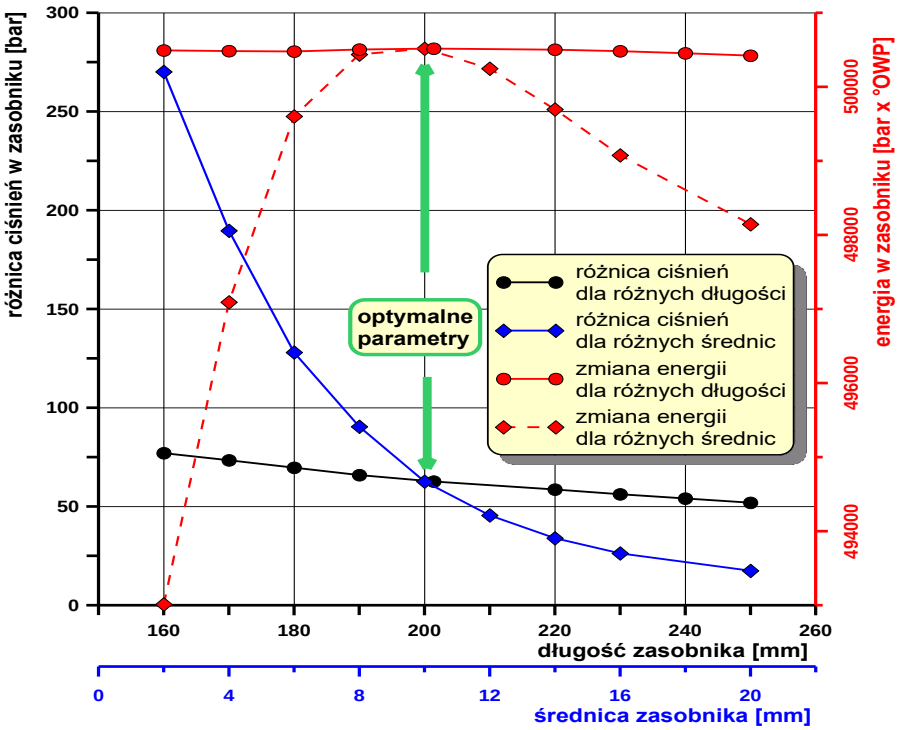
Legenda:

+ + +	wpływ zdecydowany
+	wpływ istotny
+ -	wpływ nieznaczny
- -	brak wpływu

Wspomniane wyżej różnice maksymalnych i minimalnych ciśnień w zasobniku przedstawiono na rys. 9. Różnice dla różnych długości zaznaczono czarnym kolorem i porównano z różnicami wyznaczonymi przy zmianach długości zasobnika (linia niebieska). Widać, że w miarę zwiększania długości i średnicy zasobnika zmniejszają się różnice ciśnień będące skutkiem procesu wtrysku, przy czym wpływ zmian średnicy zasobnika jest znacząco większy.

Jeśli przyjąć, że miarą będącej do dyspozycji energii paliwa przed wtryskiem jest pole pod krzywą ciśnienia w zasobniku, to ta wielkość dla różnych długości zasobnika zmienia się w stopniu równym średnim wartościom ciśnienia, czyli niewiele (rys. 9, czerwona, ciągła

linia). Na tym samym rysunku zestawiono zmiany energii paliwa w zasobniku (linie czerwonego koloru). O ile w rozważanym zakresie zmian długości nie następują istotne zmiany tej wielkości, to dla zmian średnic jest zupełnie inaczej (rys. 9, czerwona, kreskowa linia). Przebieg osiąga maksimum występuje dla zasobnika o średnicy 10 mm i długości 201,4 mm. Dla takiej konfiguracji wymiarów energia paliwa przed wtryskiem jest największa i może być właściwie wykorzystana do przygotowania mieszaniny palnej. Podane wartości jako optymalne przyjął producent analizowanego układu zasilania i stosował w zasobnikach do silników o pojemności skokowej 1700 cm^3.



Rys. 9. Obliczone różnice ciśnień i zmiany energii dla różnych wymiarów zasobnika

Należy jednak podkreślić, że zasobnik został zamodelowany w sposób uproszczony, bez uwzględniania zjawisk falowych. W rzeczywistym zasobniku paliwo podlega prawom ruchu falowego i powstają lokalne przestrzenie o ciśnieniu wyższym lub niższym w stosunku do zadanego, a te mogą w znacznym stopniu wpływać na proces dawkowania. Po ich uwzględnieniu może ulec zmianie ocena ilościowa przedstawionych zależności.

6. Podsumowanie

Proces modelowania odgrywa znaczącą rolę przy projektowaniu i doborze części maszyn. Pozwala w dużym stopniu skrócić czas wdrożenia projektowanego układu, jak również na jego dostosowanie do zabudowy równolegle z procesem projektowania. Aby model matematyczny jak najlepiej odzwierciedlał rzeczywiste zjawiska, należy sporządzić prawidłowy model fizyczny badanego układu. Oczywistym jest, że stopień uproszczenia modelowanego systemu będzie oddziaływał na dokładność wyników obliczeń, lecz w wielu przypadkach stosowanie uproszczeń jest konieczne, z uwagi na komplikację modelu matematycznego, zwiększenie czasów obliczeń, a więc obniżenie wydajności programu.

Opracowany model procesu wtrysku, dotyczący powszechnie stosowanego systemu wtrysku paliwa w silnikach o zapłonie samoczynnym typu Common Rail, pozwolił na

określenie zależności zachodzących między parametrami badanego układu. Spośród otrzymanych wyników badań, można wyodrębnić czynniki mające największy wpływ na dawkę paliwa, przebieg wtrysku, kąt początku oraz trwania wtrysku. Wpływ analizowanych wielkości na parametry wtrysku był różny, co można podsumować jak niżej.

- Przebieg ciśnienia, ma znaczący wpływ na cały proces wtrysku oraz dawkę paliwa. Zwiększenie ciśnienia w zasobniku powoduje zmianę natężenia wypływu paliwa z rozpylacza, co przekłada się na wzrost wydatku paliwa.
- Przy założonej długości, w zakresie rozważanych wartości, średnica zasobnika ma istotny wpływ na kąt początku wtrysku. Jej rosnąca wartość powoduje zwiększenie kąta początku oraz zmniejszenie ilości dawkowanego paliwa. Wynika to z objętości i ściśliwości paliwa, ponieważ większa objętość powoduje wydłużenie reakcji na sygnał wymuszający jakim jest ciśnienie paliwa w zasobniku.
- Zmienna długość zasobnika przy stałej jego średnicy ma nieznaczny wpływ na kąt początku wtrysku. Większa długości zasobnika zwiększa kąt początku wtrysku, lecz w mniejszym stopniu niż zmiana średnicy. Wynika to z mniejszego przyrostu objętości paliwa. Wielkość ta nie wpływa zarazem na kąt trwania wtrysku.
- Zmiana średnicy i długości zasobnika przy jego stałej objętości nie ma wpływu na analizowane parametry wtrysku.

Biorąc pod uwagę uzyskane wyniki symulacji, a także różnice między rezultatami obliczeń oraz wartościami zmierzonymi na stanowisku probierczym, stwierdza się dużą zgodność porównywanych wielkości. Jednak jak zawsze występują rozbieżności między układem rzeczywistym oraz modelem. Ich wartość pozwala ocenić jakość modelu oraz jego podatność na zmiany zadawanych wielkości. Mimo pewnego skomplikowania algorytmów oraz dużej ilości możliwych do zmiany parametrów jest on stosunkowo dobrze przewidywalny pod kątem generowanych wyników. Cecha ta pozwala na szybkie wykonywanie modelowych badań oraz dobór takich wartości parametrów początkowych, które umożliwiają uzyskanie żądanego przebiegu procesu wtrysku i dawkowania paliwa. Komputerowy program obliczeniowy, opracowany w oparciu o przedstawiony fizyczny model, można zakwalifikować jako dobrze odzwierciedlający badane parametry układu wtryskowego. Z uwagi na przyjęte niektóre uproszczenia w modelu matematycznym, występują różnice w wynikach obliczeń i pomiarów, lecz nie zmieniają wyników w znaczącym stopniu. Niewątpliwą niedoskonałością jest ograniczony podział dawki paliwa na części. W obecnej wersji programu można dokonać jedynie dwuczęściowego podziału. W toku dalszych prac należy dostosować program i model do aktualnych wymogów i przy wykorzystaniu wyników prac doświadczalnych opracować modele empiryczno-obliczeniowe, uwzględniające zarówno możliwość podziału na większą liczbę części, jak i większą liczbę parametrów sterujących.

References

1. Ahlin K. Modelling of pressure waves in the common rail Diesel injection system. Thesis performed in Automotive Systems at Linköping Institute of Technology, LiTH-ISY-EX-3081, 2000.

2. Amoia V, Ficarella A, Laforgia D, De Matthaeis S, Genco C. A theoretical code to simulate the behavior of an electro-injector for Diesel engines and parametric analysis. 1997, SAE Technical Paper 970349.

https://doi.org/10.4271/970349

3. Arcoumanis C, Baniasad M. S. Analysis of consecutive fuel injection rate signals obtained by the Zeuch and Bosch Methods. 1993, SAE Technical Paper 930921. https://doi.org/10.4271/930921

4. Gautier C, Sename O, Dugard L, Meissonnier G. An LFT Approach to H_ Control Design for Diesel Engine Common Rail Injection System. Oil & Gas Science and Technology, Rev. IFP, 2007,62(4): 513-522.

5. Gautier C, Sename O, Dugard L, Meissonnier G. Modelling of a Diesel Engine Common Rail Injection System. IFAC 16th Word Congress, Prague, 2005.

6. Huhtala K, Vilenius M. Study of a common rail fuel injection system. 2001, SAE Technical Paper 2001-01-3184.

7. Lino P, Maione B, Pizzo A. Nonlinear modelling and control of a common rail injection system for diesel engines. Applied Mathematical Modelling, 2007, 31(9): 1770–1784.

8. Ricco M, De Matthaeis S, Olabi A.G. Simulation of the magnetic properties for common rail electro-injector. Journal of Materials Processing Technology, 2004, 155(1): 1611–1615.

9. Schuckert M, Schultze L, Tschöke H. Zur Auslegung von Common-Rail Diesel-Einspritzsystemen. MTZ, 1998, 59(12): 800-806.

10. Seykens X.L.J, Somers L.M.T, Baer R.S.G. Detailed Modelling Of Common Rail Fuel Injection Process. Journal of Middle European Construction and Design of Cars, 2005, 3(2-3): 30-39.

11. Shinohara Y, Takeuchi K, Hermann O. E, Lumen H. J. Common-Einspritzsystem mit 3000 bar. MTZ, 2011, 72(1): 10-15.

dr inż. Jarosław Selech

Poznan University of Technology,
ul. Piotrowo 3, 60-965, Poznań, Poland,
E-mail: jaroslaw.selech@put.poznan.pl

dr hab. Karol Andrzejczak

Poznan University of Technology,
ul. Piotrowo 3A, 60-965, Poznań, Poland,
E-mail: karol.andrzejczak@put.poznan.pl

Zagregowane kryterium wyboru rozkładu czasu do uszkodzenia elementów pojazdów szynowych

Słowa klucze: *czas do uszkodzenia, estymacja rozkładu prawdopodobieństwa, niezawodność pojazdów szynowych*

Streszczenie: W pracy przedstawiono zagregowaną metodę doboru dystrybuant hipotetycznych do dystrybuanty empirycznej. Metoda miała na celu identyfikację czasu niezawodnej pracy odnawialnego obiektu technicznego poprzez zastosowanie trzech kryteriów, w których użyto następujących statystyk: zmodyfikowanej statystyki Kołmogorowa-Smirnowa (MK-S), statystyki średniego odchylenia bezwzględnego dystrybuanty hipotetycznej od empirycznej oraz statystyki obliczanej na podstawie zlogarytmowanej funkcji wiarygodności. Wartości tych statystyk posłużyły do rangowania jedenastu rozkładów prawdopodobieństwa. Dane dla których dokonano obliczeń dotyczyły uszkodzeń zamka kabiny motorniczego jakie odnotowano w ciągu pięciu lat użytkowania floty 45 tramwajów. Przed obliczeniem statystyk wyznaczono dystrybuantę empiryczną badanego elementu przy pomocy estymatora Kaplana-Meiera, a następnie przy użyciu metody największej wiarygodności oszacowano parametry uwzględnionych w badaniach rozkładów hipotetycznych. Po wyznaczaniu parametrów nastąpiło rangowanie rozkładów hipotetycznych według wartości otrzymanych dla każdego z przyjętych kryteriów, im mniejsza wartość dla danego kryterium tym wyższa pozycja w rankingu, świadcząca o lepszej jakości dopasowania według danego kryterium. Po ustaleniu rankingu według kryteriów zgodności, każdemu z kryteriów zgodności dopasowania dystrybuant modelowych do empirycznej nadano wagi. Następnie na podstawie uzyskanych trzech rankingów oraz wag nadanych poszczególnym kryteriom zgodności wyznaczana jest zagregowana miara zgodności (oznaczona DESV), która służy do wyznaczania najlepszego rozkładu prawdopodobieństwa. W prezentowanej metodzie przyjęto, że najmniejsza wartość DESV wyznacza najlepiej dopasowany rozkład hipotetyczny. W przypadku badanego elementu rozkładem tym okazał się uogólniony rozkład gamma. Pokazano, że na podstawie zagregowanego kryterium uwzględniającego trzy statystyki zgodności dopasowania zwiększa się wiarygodność estymacji rozkładu czasu pracy do uszkodzenia, unikając tym samym błędów jakie można popełnić uzależniając się tylko od jednej z nich.

1. Wprowadzenie

W tradycyjnych metodach estymacji parametrów rozkładu czasu zdatności obiektu technicznego lub jego elementu przyjmowana jest a priori określona klasa rozkładów. Celem niniejszego artykułu jest przedstawienie wyników identyfikacji najlepszych rozkładów prawdopodobieństwa czasów zdatności elementów odnawialnego obiektu technicznego z zastosowaniem zagregowanego kryterium dopasowania (aggregate matching criterion) rozkładu prawdopodobieństwa do danych empirycznych. Obiektem badań są elementy lub podzespoły aktualnie eksploatowanych jednorodnych pojazdów szynowych tworzących flotę obsługiwaną przez operatora. Dane empiryczne pozyskiwane w trakcie eksploatacji pojazdów są niepełne, gdyż w momencie zakończenia badań pojazdy są sprawne i nadal użytkowane. Tak więc autorzy niniejszego artykułu w swoich badaniach nie dysponowali pełnymi danymi o czasach zdatności wszystkich elementów

czy podzespołów badanych pojazdów szynowych. W związku z tym pojawiła się konieczność stosowania metod statystycznych uwzględniających dane cenzurowane (ucięte). Posiadając odpowiednio przygotowaną bazę danych o naprawach pojazdów badanej floty można w miarę łatwo wyznaczyć podstawowe charakterystyki niezawodnościowe wymienianych elementów [25]. Problemem natomiast okazuje się wybór dobrej miary zgodności dopasowania rozkładu czasu zdatności dla uszkadzających się elementów pojazdu. Problem ten jest przedmiotem przedstawionych w tej pracy badań i dotyczy zastosowania zagregowanego kryterium do wyznaczania najlepszych rozkładów czasów zdatności wybranych elementów pojazdu szynowego [33]. Wyniki badań zostały opracowane w postaci rankingu zgodności wybranych rodzin rozkładów z wykorzystaniem zagregowanej miary zgodności dopasowania jako kryterium.

W badaniach obiektów technicznych, jako modele czasów do ich uszkodzenia, są stosowane różne rodziny rozkładów prawdopodobieństwa [17]. Do najczęściej stosowanych w analizie czasów zdatności (Life Data Analysis LDA) należą rozkłady: normalny, wykładniczy i Weibulla [19, 10]. W przeprowadzonych badaniach poza wspomnianymi, autorzy sprawdzali możliwości użycia także innych, rzadziej stosowanych rozkładów, których jakość dopasowania do danych empirycznych dla wielu przypadków okazywała się być lepsza, od tych zazwyczaj stosowanych, są to rozkłady: logarytmiczno-normalny, gamma, uogólniony rozkład gamma, logistyczny, logarytmiczno-logistyczny oraz Gumbela [22]. Funkcje gęstości wymienionych rozkładów oraz ich parametry pokazano w tabeli 1. W przypadku uogólnionego rozkładu gamma, w celu łatwiejszej estymacji parametrów, funkcję gęstości podano także w postaci reparametryzowanej [20].

Tab. 1. Funkcje gęstości oraz parametry estymowane

Typ rozkładu	**Gęstość prawdopodobieństwa $f(t)$**	**Parametry rozkładu**
Wykładniczy	$f(t;\lambda) = \lambda e^{-\lambda t},\ t \geq 0, \lambda > 0$	$\frac{1}{\lambda}$ – parametr skali
Dwuparametrowy wykładniczy	$f(t;\lambda,\gamma) = \lambda e^{-\lambda(t-\gamma)},\ t \geq \gamma, \lambda > 0$	$\frac{1}{\lambda}$ – parametr skali γ – parametr położenia
Normlany	$f(t;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2},\ t \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0$	μ – wartość oczekiwana σ – odchylenie standardowe
Logarytmo-normalny	$f(t;\mu',\sigma') = \frac{1}{t\cdot\sigma'\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t)-\mu'}{\sigma'}\right)^2},\ t > 0, \mu' \in \mathbb{R}, \sigma' > 0$	μ' – wartość oczekiwana $\ln T$ o rozkładzie normalnym, σ' – odchylenie standardowe $\ln T$
Dwuparametrowy Weibulla	$f(t;\beta,\eta) = \frac{\beta}{\eta}\cdot\left(\frac{t}{\eta}\right)^{\beta-1}\cdot e^{-\left(\frac{t}{\eta}\right)^{\beta}},\ t \geq 0, \beta > 0, \eta > 0$	η – parametr skali β – parametr kształtu
Trójparametrowy Weibulla	$f(t;\beta,\eta,\gamma) = \frac{\beta}{\eta}\left(\frac{t-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{t-\gamma}{\eta}\right)^{\beta}},\ t \geq \gamma, \beta > 0, \eta > 0, \gamma \in \mathbb{R}$	η – parametr skali β – parametr kształtu γ – parametr położenia
Gamma	$f(t;\mu,\kappa) = \frac{\exp(\kappa\,(\ln(t)-\mu)-\exp(\ln(t)-\mu))}{t\,\Gamma(\kappa)};\ t > 0, \mu \in \mathbb{R}, \kappa > 0$	e^{μ} – parametr skali κ – parametr kształtu
Uogólniony Gamma	$f(t;\theta,\beta,\kappa) = \frac{\beta}{\Gamma(\kappa)\cdot\theta}\cdot\left(\frac{t}{\theta}\right)^{\kappa\beta-1}\cdot e^{-\left(\frac{t}{\theta}\right)^{\beta}},\ \theta > 0,\ \beta > 0, \kappa > 0$ Reparametryzacja: $\mu = \ln(\theta) + \frac{1}{\beta}\ln\left(\frac{1}{\lambda^2}\right);\ \sigma = \frac{1}{\beta\sqrt{\kappa}};\ \lambda = \frac{1}{\sqrt{\kappa}}$ $f(t;\mu,\sigma,\lambda) = \begin{cases} \frac{\lvert\lambda\rvert}{\sigma\cdot t}\cdot\frac{1}{\Gamma\left(\frac{1}{\lambda^2}\right)}\cdot\exp\left[\frac{\lambda\frac{\ln(t)-\mu}{\sigma}+\ln\left(\frac{1}{\lambda^2}\right)-\exp\left(\lambda\frac{\ln(t)-\mu}{\sigma}\right)}{\lambda^2}\right] & \text{dla } \lambda > 0 \\ \frac{1}{t\cdot\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(t)-\mu}{\sigma}\right)^2\right) & \text{dla } \lambda = 0 \end{cases}$ $t \geq 0, \mu \in \mathbb{R}, \lambda \geq 0, \sigma > 0,$	θ – parametr skali, β – parametr kształtu κ – parametr kształtu

Logistyczny	$f(t;\mu,\sigma) = \frac{\exp\left(\frac{t-\mu}{\sigma}\right)}{\sigma\left(1+\exp\left(\frac{t-\mu}{\sigma}\right)\right)^2}, t \in \mathbb{R},\ \mu \in \mathbb{R}, \sigma > 0$	σ – parametr skali. μ – parametr położenia
Logarytmo-logistyczny	$f(t;\mu,\sigma) = \frac{\exp\left(\frac{\ln(t)-\mu}{\sigma}\right)}{\sigma t\left(1+\exp\left(\frac{\ln(t)-\mu}{\sigma}\right)\right)^2},\ t > 0,\ \mu \in \mathbb{R}, \sigma > 0$	μ – parametr skali σ – parametr kształtu
Gumbela	$f(t;\mu,\sigma) = \frac{1}{\sigma}\exp\left(\left(\frac{t-\mu}{\sigma}\right) - \exp\left(\frac{t-\mu}{\sigma}\right)\right), \sigma > 0$	μ – parametr położenia σ – parametr skali

Estymację parametrów tych rozkładów można przeprowadzać metodami analitycznymi, numerycznie oraz graficznie [16, 26, 29]. Do najczęściej stosowanych metod zalicza się metodę momentów, metodę największej wiarygodności, metodę najmniejszych kwadratów, metodę dopasowania w siatkach rozkładów oraz metodę współczynnika korelacji wykresu prawdopodobieństwa (ang. probability plot correlation coefficient (PPCC)) [1, 38, 32]. W praktyce inżynierskiej najczęściej stosuje się metody numeryczne i graficzne z użyciem specjalistycznych narzędzi informatycznych [12, 39]. Na podstawie danych eksploatacyjnych i wyboru metody estymacji dokonywana jest estymacja parametrów (kształt, skala, położenie) dla wybranych rodzin rozkładów prawdopodobieństwa [28, 37]. Mając oszacowane różne rozkłady można wskazać wśród nich najlepiej dopasowany do danych empirycznych w sensie najmniejszej sumy kwadratów odchyleń.

W proponowanej metodyce identyfikacji czasu zdatności wybranego elementu pojazdu uwzględnione są wszystkie dostępne dane dotyczące czasów/przebiegów pojazdów do uszkodzenia tego elementu we wszystkich pojazdach badanej floty. Obejmuje to przypadek, w którym dany element jest zdatny w momencie przerwania badań, a czas zdatności takiego elementu nazywany jest prawostronnie cenzurowanym. Sposób przygotowania danych statystycznych na podstawie bazy danych eksploatacyjnych opracowany został w artykułach [3, 2].

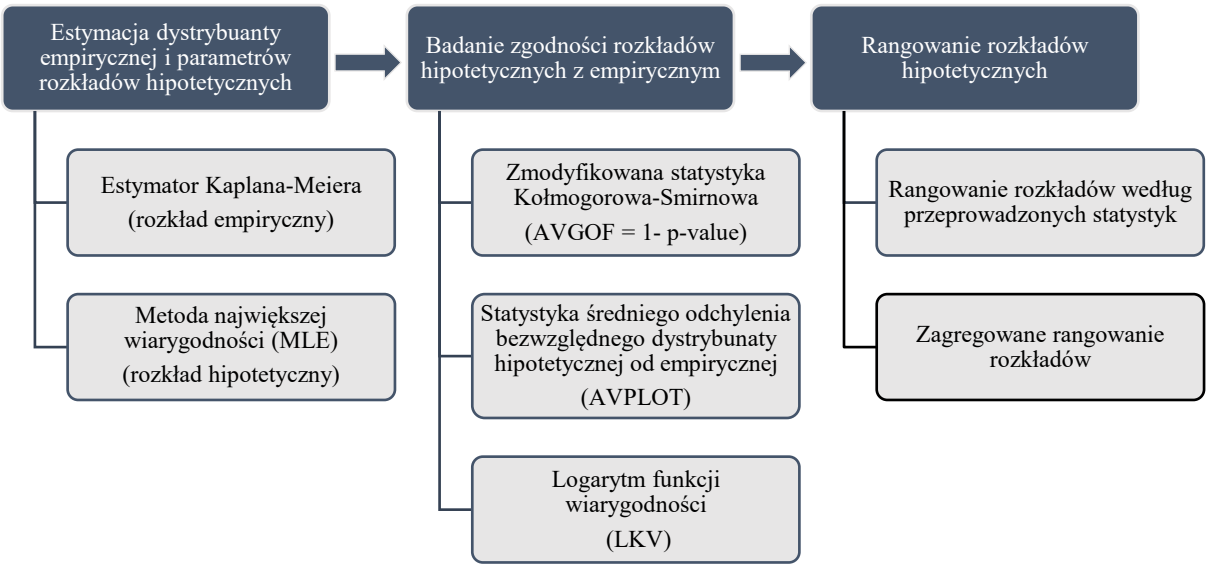
Zamiast tradycyjnego jednokryterialnego wyboru najlepiej dopasowanej rodziny rozkładów prawdopodobieństwa autorzy proponują zastosowanie zagregowanego kryterium uwzględniającego trzy miary zgodności dopasowania rozkładów hipotetycznych. W kryterium tym brany jest pod uwagę ranking zgodności dopasowania poszczególnych modeli probabilistycznych do danych empirycznych w tym do prawostronnie cenzurowanych danych eksploatacyjnych badanej floty pojazdów.

W zagregowanej metodzie wybór rozkładu poprzedzony jest rankingiem rozkładów dla trzech kryteriów zgodności. Estymacji parametrów wybranych rodzin rozkładów dokonano za pomocą programu Weibull++ używając modułu Distribution Wizard, który po przeprowadzeniu odpowiednich obliczeń szereguje rozkłady począwszy od najbardziej dopasowanych. Zanim jednak zostanie przeprowadzone badanie zgodności rozkładów, wyznaczana jest dystrybuanta lub funkcja niezawodności rozkładu empirycznego metodą Kaplana-Meiera, a następnie wyznaczane są parametry rozkładów hipotetycznych metodą największej wiarygodności (MLE – *Maximum Likelihood Estimation*).

Kolejny krok polega na wyznaczeniu statystyk zgodności dopasowania poszczególnych dystrybuant hipotetycznych do dystrybuanty empirycznej oznaczonej F_n. Na tej podstawie dokonywany jest ranking jedenastu rozkładów zestawionych w tabeli 1 stosowanych w analizie przeżycia [18, 30]. Jeśli tylko założenia są spełnione, to rankingi zgodności rozkładów przeprowadzane są niezależnie według trzech kryteriów, z użyciem zmodyfikowanej statystyki Kołmogorowa-Smirnowa (MK-S), średniego odchylenia bezwzględnego dystrybuanty hipotetycznej od empirycznej oraz wartości zlogarytmowanej funkcji wiarygodności [23].

W ostatecznym ustaleniu rankingu rozkładów brane są pod uwagę rankingi uzyskane przez te trzy kryteria z uwzględnieniem wag przypisanych do każdego z nich. Po nadaniu wag kryteriom, wyliczana jest wartość końcowa DESV (ang. *Distribution Estimation Values*), która wskazuje najlepiej dopasowany rozkład wyznaczony według opisanego zagregowanego kryterium. Schemat kolejnych kroków obliczeniowych w zagregowanej metodzie rangowania rozkładów ukazano na rys. 1.

Zgodnie z tym schematem w pierwszym kroku, opierając się na pozyskanych danych oraz analizując długość czasu obserwacji (uciętej prawostronnie) oszacowano estymatorem Kaplana-Meiera parametry funkcji przeżycia i wyznaczono dystrybuantę empiryczną [7]. Następnie w celu wyznaczenia parametrów jedenastu rozkładów hipotetycznych, zestawionych w tabeli 2 zastosowano metodę największej wiarygodności [15, 11].



Rys. 1. Schemat zagregowanego kryterium rankingowania rozkładów

W drugim kroku, kolejno dla jedenastu rozkładów, za pomocą statystyk zgodności oceniana jest hipoteza zerowa

$$H_0\colon\ T \sim F \tag{1}$$

orzekająca, że czas do uszkodzenia T badanego elementu pojazdu ma rozkład prawdopodobieństwa o dystrybuancie F z wyznaczonymi parametrami. Ocena ta jest dokonywana na podstawie próby losowej $T_1, T_2, \dots, T_n$ dotyczącej czasów do uszkodzenia badanego elementu. W niniejszej pracy czasy do uszkodzenia badanego elementu wyrażone są przebiegiem liczonym w kilometrach podobnie jak w pracy [2].

2. Kryteria rankingowania rozkładów hipotetycznych

Wśród zastosowanych kryteriów zgodności szczególną rolę pełni zmodyfikowana statystyka Kołmogorowa-Smirnowa (AVGOF – *average goodness of fit*), która ocenia różnicę statystyczną pomiędzy wartościami dystrybuanty empirycznej i hipotetycznej. Ta szczególna rola tej statystyki wynika z faktu jej dużej wrażliwości na lokalne odchylenia. Ponadto można ją stosować nawet przy małej liczbie danych oraz przy nieznanych parametrach rozkładu hipotetycznego. Zastosowanie statystyki MK-S jest więc koniecznością, gdy parametry badanych rozkładów muszą być estymowane.

Ponieważ rozkłady statystyk MK-S zależą od hipotetycznej rodziny rozkładów, której parametry są estymowane, więc dla każdego rozkładu wyznaczana jest wartość krytyczna, przy której odrzucana jest hipoteza zerowa [30]. Analityczne wyznaczenie wartości krytycznej często jest trudne lub wręcz niemożliwe, dlatego do jej uzyskania stosowana jest metoda Monte Carlo [6, 18].

Statystka MK-S użyta do sprawdzenia zgodności rozkładu hipotetycznego z rozkładem empirycznym korzysta ze statystyki D_{max} określonej jako maksimum bezwzględnej różnicy między

wartością dystrybuanty empirycznej $F_n(t)$, a dopasowaną dystrybuantą hipotetyczną $F(t)$ i określona jest wzorem [18]:

$$D_{max} = \max_{1 \le i \le n} |F_n(t_i) - F(t_i)| \tag{2}$$

gdzie:

D_{max} – wartość statystyki,
n – liczebność próby,
$F_n(t_i)$ – wartość dystrybuanty empirycznej,
$F(t_i)$ – wartość dystrybuanty hipotetycznej.

Natomiast wartość krytyczna D_{CRIT} w zmodyfikowanej statystyce Kołmogorowa-Smirnowa jak już wspomniano, ze względu na uciążliwość obliczeń, wyznaczana jest metodą Monte-Carlo.

Statystyka MK-S służy do ustalenia prawdopodobieństwa odrzucenia hipotezy zerowej, tj. prawdopodobieństwa zdarzenia $D_{CRIT} < D_{max}$. Tak więc w pierwszym kryterium podstawą porządkowania hipotetycznych rozkładów jest prawdopodobieństwo:

$$\mathrm{P}(D_{CRIT} < D_{max}) \tag{3}$$

Im większa wartość statystyki D_{max} tym coraz bardziej istotna jest różnica pomiędzy rozkładem hipotetycznym określonym przez dystrybuantę F a rozkładem empirycznym określonym przez dystrybuantę F_n. Ponieważ wartość krytyczna D_{CRIT} wyznaczana jest metodą Monte Carlo poprzez m-krotne generowanie n chwil uszkodzeń $t_{s1}, t_{s2}, \dots t_{sn}$ dla których tworzone są realizacje dystrybuant symulacyjnych $F_s(t_{si}), s = 1, 2, \dots, m$ i dla każdej z nich wyznaczane są maksymalne różnice z wartościami dystrybuanty hipotetycznej

$$d_{max,s} = \max_{1 \le s \le m} |F_s(t_{si}) - F(t_{si})|, \qquad s = 1, 2, \dots, m \tag{4}$$

więc do oszacowania wartości krytycznej D_{CRIT} przyjęta jest średnia arytmetyczna d_{CRIT} określona wzorem [6]:

$$\widehat{D_{CRIT}} = d_{CRIT} = \frac{1}{m} \sum_{s=1}^{m} d_{max,s} \tag{5}$$

Ostatecznie w kryterium MK-S do oceny zgodności rozkładów przyjmujemy;

$$\mathrm{AVGOF}(F_n, F) = 100 \cdot \mathrm{P}(d_{CRIT} < D_{max}) \tag{6}$$

Duże wartości AVGOF, zbliżone do 100, wskazują, że istnieje istotna różnica między rozkładem hipotetycznym a danymi empirycznymi. Stąd rozkład hipotetyczny jest tym lepszy im mniejsza jest wartość statystyki AVGOF.

W drugim kryterium zgodności badane jest średnie odchylenie bezwzględne dystrybuanty hipotetycznej od empirycznej, a statystyka użyta do oceny zgodności oznaczana AVPLOT (*average plot fit*) wyznaczana jest wg wzoru:

$$\mathrm{AVPLOT}(F_n, F) = 100\frac{1}{n}\sum_{i=1}^{n}|F_n(t_i) - F(t_i)| \tag{7}$$

gdzie:
n – liczebność próby,
$F_n(t_i)$ – wartości dystrybuanty empirycznej,
$F(t_i)$ – wartości dystrybuanty hipotetycznej.

Kryterium to w przeciwieństwie do kryterium MK-S nie jest wrażliwe na lokalne odchylenia, uwzględnia natomiast globalne zróżnicowanie rozkładów i stanowi dobre uzupełnienie do kryterium MK-S.

W trzecim kryterium badania zgodności rozkładów użyta została funkcja wiarygodności (LKV – *Likelihood Value Test*), jako miara dopasowania modelu probabilistycznego do danych empirycznych. Obliczana jest wartość logarytmu funkcji wiarygodności (LKV) dla danych empirycznych [27, 14]. Funkcja wiarygodności L zależy od realizacji próby losowej $T_1, T_2, \dots, T_n$ oraz od parametrów θ_j, dla których przyjmuje ona maksymalne wartości. Ogólna postać funkcji wiarygodności przedstawiona jest zależnością [33, 30]:

$$L\left(\theta_1, \theta_2, \dots, \theta_k \middle| T_1, T_{2,\dots,} T_n\right) = \prod_{i=1}^{n} f(T_i; \theta_1, \theta_2, \dots, \theta_k) \tag{8}$$

gdzie:
n – liczba uszkodzonych elementów,
k – liczba parametrów,
$\theta_j, j = 1, 2, \dots, k$ – parametry rozkładu,
$T_i, i = 1, 2, \dots, n$ – czas do uszkodzenia *i*-tego elementu,

W badanym przypadku, funkcja została rozszerzona o czynniki uwzględniające dane ucięte prawostronnie. Zlogarytmowana funkcja wiarygodności jest sumą logarytmów gęstości prawdopodobieństwa dla poszczególnych czasów życia badanego elementu [18, 30]:

$$\Lambda\,(\theta_1, \theta_2, \dots, \theta_k) = \ln L\left(\theta_1, \theta_2, \dots, \theta_k \middle| T_1, T_{2,\dots,} T_n\right) = \sum_{i=1}^{n} \ln f(T_i; \theta_1, \theta_2, \dots, \theta_k) \tag{9}$$

gdzie:
L – funkcja wiarygodności,
n – liczba uszkodzonych elementów,
$\theta_j,\ j = 1, 2, \dots, k$ – parametry rozkładu,
$T_i,\ i = 1, 2, \dots, n$ – czas do uszkodzenia *i*-tego elementu,

Wartości estymatorów nieznanych parametrów $\theta_1, \theta_2, \dots, \theta_k$ są wyznaczane przez maksymalizację zlogarytmowanej funkcji wiarygodności $\Lambda\,(\theta_1, \theta_2, \dots, \theta_k)$. Warunkiem koniecznym istnienia ekstremum tej funkcji jest przyjmowanie wartości 0 przez jej wszystkie pochodne cząstkowe.

W celu wyznaczenia estymatorów nieznanych parametrów wyznaczane są pochodne cząstkowe $\frac{\partial\Lambda(\theta_1,\theta_2,\dots,\theta_k)}{\partial\theta_j}$ funkcji Λ względem parametrów θ_j, $j = 1, 2, \dots, k$. Aby oszacować parametry należy każdą pochodną cząstkową przyrównać do zera i rozwiązać k równań:

$$\frac{\partial\Lambda(\theta_1, \theta_2, \dots, \theta_k)}{\partial\theta_1} = 0$$

$$\dots\dots\dots$$

$$\frac{\partial \Lambda(\theta_1, \theta_2, \dots, \theta_k)}{\partial \theta_k} = 0 \tag{10}$$

W ostatnim kroku, na podstawie każdego z trzech kryteriów zgodności dla wszystkich 11 rozkładów nadawane są rangi od najlepiej do najgorzej dopasowanego rozkładu hipotetycznego do danych empirycznych. Na tej podstawie dla każdego kryterium zgodności z osobna porządkowane są rozkłady hipotetyczne poprzez przypisanie im kolejnych liczb naturalnych. Na koniec na podstawie uzyskanych trzech rankingów oraz wag nadanych poszczególnym kryteriom zgodności wyznaczana jest zagregowana miara zgodności DESV. Miara ta dla i-tej dystrybuanty hipotetycznej F_i wyznaczana jest ze wzoru (11):

$$\mathrm{DESV}(F_i) = \mathrm{RAVGOF}(F_i) \cdot \mathrm{WAVGOF} + \mathrm{RAVPLOT}(F_i) \cdot \mathrm{WAVPLOT} + \mathrm{RLKV}(F_i) \cdot \mathrm{WLKV} \tag{11}$$

gdzie:
$\mathrm{RAVGOF}(F_i)$ oznacza rangę rozkładu F_i według kryterium AVGOF,
$\mathrm{RAVPLOT}(F_i)$ oznacza rangę rozkładu F_i według kryterium AVPLOT,
$\mathrm{RLKV}(F_i)$oznacza rangę rozkładu F_i według kryterium LKV,
WAVGOF oznacza wagę kryterium AVGOF,
WAVPLOToznacza wagę kryterium AVPLOT,
WLKV oznacza wagę kryterium LKV.

Zagregowane kryterium DESV zgodności dopasowania rozkładów jest więc średnią ważoną poszczególnych rang rozkładów hipotetycznych. Po wyliczeniu wartości DESV dla poszczególnych hipotetycznych rozkładów wyznaczany jest ich ostateczny ranking. Rozkład, który uzyska najmniejszą wartość DESV jest pozycjonowany jako najlepiej dopasowany według zaprezentowanego zagregowanego kryterium zgodności i otrzymuje w rankingu numer 1. Zagregowane kryterium jest zastosowane do podjęcia ostatecznej decyzji wyboru najlepiej dopasowanego rozkładu hipotetycznego do danych empirycznych spośród badanych rozkładów hipotetycznych.

3. Przedmiot badań

Zagregowane kryterium rangowania rozkładów czasów zdatności wybranych elementów pojazdu przeprowadzono na podstawie danych eksploatacyjnych floty 45 jednorodnych pojazdów szynowych transportu miejskiego. Badano pięcioczłonowe niskopodłogowe tramwaje Tramino S105P o całkowitej masie 42,5 tony i około 32 metrach długości. Są to wagony przegubowe i jednoprzestrzenne. Tramwaj może zabrać na pokład maksymalnie 229 pasażerów, w tym 48 na miejscach siedzących. Zakres danych eksploatacyjnych obejmował pięć początkowych lat użytkowania floty, w tym dwa lata objęte gwarancją, a trzy kolejne kontraktem serwisowym [31, 9]. Wszystkie tramwaje użytkowane były w podobnych warunkach eksploatacyjnych tj. tej samej infrastrukturze torowej, zbliżonych dziennych i rocznych przebiegach oraz takim samym harmonogramem i zakresem serwisowania (obsług prewencyjnych).

Rys. 2. Tramwaj Tramino S105P

Z bazy danych o uszkodzeniach tramwajów badanej floty, do badania rozkładu czasu do uszkodzenia wybrano zamek otwierania drzwi kabiny motorniczego. Uszkodził się on 54 razy w ciągu pierwszych pięciu lat eksploatacji badanej floty i wygenerował koszty w wysokości 0,52% wszystkich obsług korekcyjnych [5]. Zamek zamocowany jest na drzwiach pomiędzy przestrzenią pasażerską a kabiną motorniczego. Aby otworzyć drzwi kabiny motorniczego od zewnątrz należy specjalnym kluczem odblokować mechanicznie zamek. Uszkodzeniu ulegał element ryglujący zamek, blokując się i tym samym uniemożliwiając otwarcie drzwi i wejście do pojazdu motorniczemu. Należało wówczas w zależności od miejsca, w którym nastąpiło uszkodzenie, wezwać serwis lub siłowo wyważyć drzwi, uszkadzając konstrukcję mocowania zaczepu rygla. Za każdym razem, uszkodzony zamek wymieniany był na nowy. Przyczyną uszkadzania zamka było zbyt intensywne zużywanie się mechanizmu wewnętrznego odpowiedzialnego za wysuwanie rygla, spowodowane źle dobranym materiałem konstrukcyjnym, w wyniku czego zamek zacinał się, a czasem uniemożliwiał wyjęcie włożonego do niego klucza. Fotografię zamka zamieszczono na rys. 3.

Rys. 3. Zamek otwierania drzwi kabiny motorniczego

4. Dane empiryczne

Proces eksploatacji tramwajów stanowi cenne źródło informacji konieczne do oceny niezbędnych charakterystyk niezawodnościowych oraz prognozowania kosztów ich utrzymania. Pod pojęciem informacji eksploatacyjnej należy rozumieć wszelkie dane o zdarzeniach zachodzących w fazie użytkowania i obsługiwania tramwajów [13]. Dane te odgrywają kluczową rolę w planowaniu i bieżącym zarządzaniu eksploatacją floty pojazdów, jak również w doskonaleniu technologii i konstrukcji pojazdów [4, 35]. Informacje eksploatacyjne odgrywają szczególnie ważną rolę dla przewoźników użytkujących pojazdy, gdyż umożliwiają właściwe planowanie kosztów eksploatacji, przeglądów i napraw oraz ocenę stosowania środków transportu [24, 34].

Przed przystąpieniem do estymacji parametrów probabilistycznych modeli czasów zdatności wybranych elementów pojazdu należy odpowiednio przygotować dane eksploatacyjne. Dla badanej floty użytkowanych tramwajów dane eksploatacyjne dotyczące poszczególnych elementów pojazdu są ucięte prawostronnie typu I, co oznacza że dla ustalonego czasu użytkowania floty tramwajów tylko w części pojazdów zamek uległ uszkodzeniu i był wymieniany na nowy, a w niektórych pojazdach był wymieniany wielokrotnie. Ponieważ obiektem badań są intensywnie użytkowane pojazdy, więc czasy do uszkodzenia poszczególnych elementów tramwaju liczone są przebiegiem wyrażonym w kilometrach. Znany jest czas rozpoczęcia eksploatacji każdego pojazdu oraz rejestrowane są przebiegi tramwajów przy których nastąpiły uszkodzenia jego elementów [33, 2]. Z przebiegów tramwajów wyznaczane są przebiegi uszkadzających się elementów. Metodę wyznaczania przebiegu uszkadzających się elementów pojazdu przedstawiono w pracy [3].

Odpowiednio przygotowane dane zestawiono w tab. 2. Zawierają one dokładny czas do chwili uszkodzenia badanego elementu – zamka kabiny motorniczego z floty 45 badanych tramwajów wyrażony przebiegiem w kilometrach i oznaczony jako F (ang. *failure*) oraz czas pracy pozostałych nieuszkodzonych zamków, oznaczonych jako S (ang. *suspension*), także wyrażony osiągniętym przebiegiem do chwili zatrzymania badań. W momencie zatrzymania badań wszystkie zamki w 45 pojazdach były sprawne, mimo że wiele z nich było wymienianych z powodu uszkodzenia. Ponieważ głównym powodem wymiany zamka na nowy jest uszkadzanie mechanizmu otwierania, dlatego uszkodzenia tego typu zakwalifikowano do kategorii uszkodzeń mechanicznych.

Tab. 2. Prawostronnie cenzurowane czasy do uszkodzenia zamka w 5-letnim okresie eksploatacji

Przebieg [km]	F/S	Przebieg [km]	F/S	Przebieg [km]	F/S	Przebieg [km]	F/S
174 124	F	256 382	F	114 128	S	67 733	F
196 837	S	144 819	S	135 078	F	300 557	S
317 275	S	223 684	F	136 600	F	103 378	F
292 525	F	46 217	F	97 832	S	177 506	F
112 431	S	43 897	S	377 101	S	23 153	S
196 218	F	155 522	F	93 585	F	242 544	F
1 910	F	201 423	S	238 103	S	89 047	S
93 529	S	119 376	F	285 538	F	125 785	F
334 484	S	198 190	S	43 117	S	58 407	F
366 935	F	368 449	S	221 226	F	117 646	S
28 826	S	340 330	F	117 701	S	202 396	F
191 367	F	58 964	S	28 934	F	127 143	S
21 117	F	193 641	F	135 673	F	287 695	F
135 831	S	155 920	S	155 828	S	53 863	S
348 956	F	206 246	F	92 594	F	174 580	F
38 020	S	144 352	S	197 981	S	139 571	S
188 493	F	371 800	S	148 840	F	210 775	F
70 534	F	22 482	F	27 858	F	102 038	S
102 343	S	139 974	F	107 491	F	131 537	F
340 236	F	39 840	F	52 280	S	126 738	F
52 022	S	127 333	F	250 370	S	83 497	S
115 592	F	21 021	S	282 989	F	176 928	F
79 071	F	376 601	S	77 834	S	81 021	F

72 135	F	354 513	S	86 028	F	103 807	S
105 552	S	203 105	F	226 082	S	-	-

F – uszkodzenie, S – przeżycie

Na podstawie danych z tabeli 2 dokonano estymacji parametrów dla 11 hipotetycznych rozkładów. Wyniki estymacji wszystkich badanych rozkładów podano w zestawieniu tabelarycznym (tab. 3).

Tab. 3. Oszacowane parametry badanych rozkładów

1P-Exponential $\hat{\lambda}$ = 3,413E-06	**2P-Exponential** $\hat{\lambda}$ = 3,937E-06 $\hat{\gamma}$ = 21117	**Normal** $\hat{\mu}$ = 218 279,5 $\hat{\sigma}$ = 115 461,8	**Lognormal** $\hat{\mu}'$ = 12,184 $\hat{\sigma}'$ = 0,819
2P-Weibull $\hat{\beta}$ = 1,745 $\hat{\eta}$ = 255 316,9	**3P-Weibull** $\hat{\beta}$ = 1,885 $\hat{\eta}$ = 266 209,6 $\hat{\gamma}$ = -10 300,12	**Gamma** $\hat{\mu}$ = 11,53 $\hat{\kappa}$ = 2,307	**G-Gamma** $\hat{\mu}$ = 12,415 $\hat{\sigma}$ = 0,605 $\hat{\lambda}$ = 0,857
Logistic $\hat{\mu}$ = 211 755,9 $\hat{\sigma}$ = 68 491,7	**Loglogistic** $\hat{\mu}$ = 12,198 $\hat{\sigma}$ = 0,447	**Gumbel** $\hat{\mu}$ = 274 770,6 $\hat{\sigma}$ = 104 341,5	

5. Identyfikacja najlepszego rozkładu prawdopodobieństwa

Do wyboru najlepszego rozkładu hipotetycznego spośród 11 rozważanych zastosowano opisane w części 2 zagregowane kryterium ich rangowania. W ustaleniu rankingu rozkładów przeprowadzono najpierw estymację parametrów hipotetycznych rozkładów, a następnie przeprowadzono ranking tychże rozkładów na podstawie opisanych trzech kryteriów. Wyniki tego rankingu zestawione są w tab. 4.

Pierwsza kolumna wskazuje nazwę rozkładu prawdopodobieństwa. Druga zawiera wartości statystyki Kołmogorowa-Smirnowa AVGOF – prawdopodobieństwo odrzucenia hipotezy roboczej dla statystki MK-S. Kolumna trzecia opisana jako AVPLOT przedstawia wartość średnią bezwzględnych wartości różnic pomiędzy dystrybuantami empiryczną a hipotetyczną. Czwarta kolumna oznaczona jako LKV przedstawia miary dopasowania rozkładów wyznaczone z zastosowaniem kryterium zlogarytmowanej funkcji wiarygodności [8, 36, 21].

Tab. 4. Wyniki poszczególnych statystyk dla danych z tabeli 1

Rozkład	AVGOF	AVPLOT	LKV
1P-Exponential	80,740	7,599	-720,16
2P-Exponential	55,253	5,666	-712,58
Normal	30,946	4,178	-715,65
Lognormal	14,276	2,790	-711,82
2P-Weibull	1,034	1,709	-709,71
3P-Weibull	2,396	1,870	-710,23
Gamma	0,045	1,585	-709,78
G-Gamma	**0,289**	**1,589**	**-709,66**
Logistic	36,386	3,730	-717,34
Loglogistic	1,716	1,768	-710,49
Gumbel	84,250	6,355	-723,73

Po wyznaczeniu statystyk zgodności dla trzech kryteriów i nadaniu rang poszczególnym rozkładom prawdopodobieństwa, kolejnym krokiem było przypisanie wag trzem kryteriom zgodności. W pracy użyto wartości wag dobranych przez producenta oprogramowania jako domyślne, wyznaczone na podstawie praktyki inżynierskiej, wynikającej z wielu analiz prowadzonych w zastosowaniach przemysłowych. Stosując wagi przypisane każdemu z kryteriów obliczono średnią ważoną dla uzyskanych rang poszczególnych kryteriów. Na koniec za pomocą

zaprezentowanego zagregowanego kryterium DESV dokonano ostatecznego rankingu jedenastu hipotetycznych rozkładów. W estymacji parametrów hipotetycznych rozkładów oraz w ich rankingowaniu korzystano z programu Weibull++. Dla analizowanych danych przyjęto następujące wagi dla poszczególnych kryteriów AVGOF – 40, AVPLOT – 10 i LKV – 50. Po wyliczeniu wartości DESV wyznaczano ranking rozkładów (tab. 5). Rozkład, który uzyskał najmniejszą wartość DESV jest pozycjonowany jako najlepiej dopasowany według zaprezentowanego zagregowanego kryterium i otrzymał w rankingu numer 1. Jak widać z tab. 5 najmniejszą wartości statystyki DESV uzyskano dla uogólnionego rozkładu gamma. Obliczona ona została ze wzoru (11) w następujący sposób:

$$\mathrm{DESV} = (2 \times 40) + (2 \times 10) + (1 \times 50) = 150 \tag{12}$$

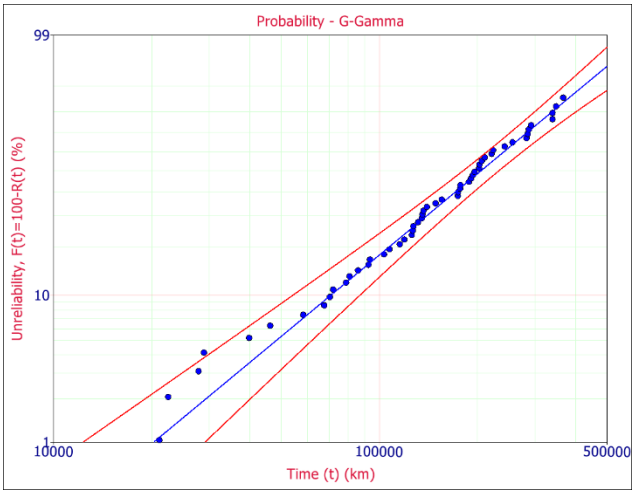
Tak więc, dla danych zawartych w tabeli 2, dotyczących uszkodzeń zamka w ciągu 5 lat eksploatacji floty tramwajów, stosując opracowane zagregowane kryterium jako najlepszy wybrano uogólniony rozkład gamma, co jest odnotowane w ostatniej kolumnie tab. 5.

Tab. 5. Wartości średnich ważonych oraz ranking rozkładów

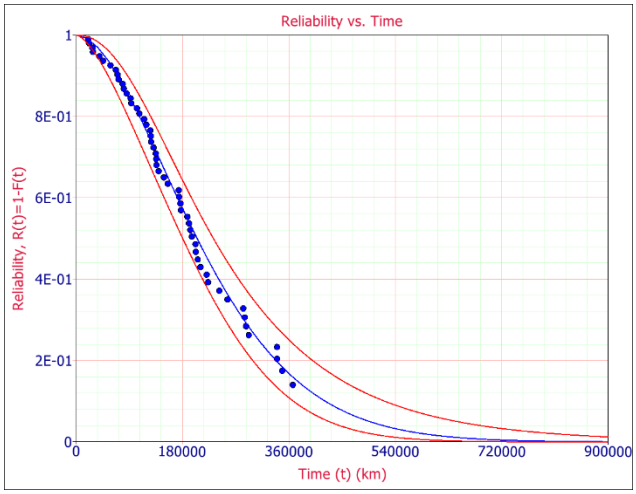
Rozkład	AVGOF	AVPLOT	LKV	DESV	Ranking
1P-Exponential	10	11	10	1010	9
2P-Exponential	9	9	7	800	7
Normal	7	8	8	760	6
Lognormal	6	6	6	600	5
2P-Weibull	3	3	2	250	3
3P-Weibull	5	5	4	450	4
Gamma	1	1	3	200	2
G-Gamma	**2**	**2**	**1**	**150**	**1**
Logistic	8	7	9	840	8
Loglogistic	4	4	5	450	4
Gumbel	11	10	11	1090	10

Oszacowane reparametryzowane parametry μ, σ, λ dla tego rozkładu uzyskały następujące wartości $\hat{\mu} = 12{,}415$; $\hat{\sigma} = 0{,}6058$; $\hat{\lambda} = 0{,}8572$. Obliczona intensywność uszkodzeń zamka wyniosła 0,000000617/km, a średni czas do uszkodzenia 229 623 km.

Dla zobrazowania dopasowania wytypowanego rozkładu na rys. 4 zaprezentowano dane na siatce uogólnionego rozkładu gamma. Na kolejnych rysunkach przedstawiono funkcję niezawodności (rys. 5), gęstości prawdopodobieństwa (rys. 6) oraz histogram liczebności uszkodzeń (rys. 7).



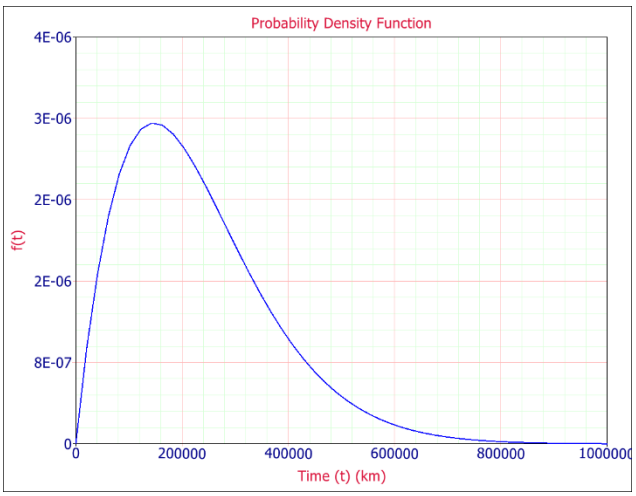
Rys. 4. Przedstawienie danych w siatce probabilistycznej uogólnionego rozkładu gamma



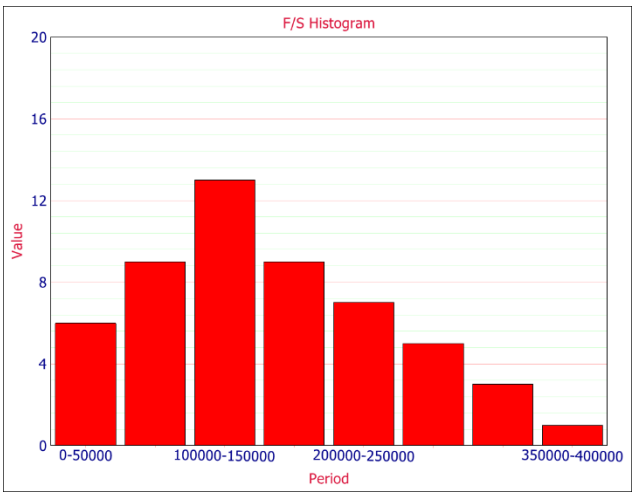
Rys. 5. Funkcja niezawodności

Na rys. 4. niebieska linia reprezentuje zamodelowane prawdopodobieństwo wystąpienia uszkodzeń zgodnie z uogólnionym rozkładem gama, a czerwona linia to dwustronne 95% przedziały ufności. Wykres funkcji niezawodności (rys. 5) pokazuje zmianę wartości niezawodności w czasie wyrażonym jako przebieg w km, wskazując trend w zachowaniu się badanego elementu w przypadku kolejnych uszkodzeń. Wykres funkcji gęstości prawdopodobieństwa uszkodzeń, umożliwia wizualizację rozkładu danych w czasie (rys. 6). Wykres histogramu (rys. 7) pokazuje, że stosunkowo duża część uszkodzeń mieści się między wartościami 50 000 a 200 000 km.

Prezentacja graficzna estymowanych charakterystyk funkcyjnych (niezawodność, gęstość prawdopodobieństwa) oraz histogramu liczebności uszkodzeń może zostać użyta w celu łatwiejszego ustalenia postaci uszkodzenia. Informacja ta jest istotna w przypadku prognozowania przebiegu uszkodzeń oraz wyznaczania kosztów obsług korekcyjnych z nimi związanych w przyszłości.



Rys. 6. Funkcja gęstości prawdopodobieństwa



Rys. 7. Histogram

Z przedstawionych badań czasu do uszkodzenia zamka kabiny motorniczego wynika, że najlepiej dopasowanym rozkładem wyznaczonym według zagregowanego kryterium jest uogólniony rozkład gamma. Należy również zaznaczyć, iż wraz z pojawiającymi się kolejnymi uszkodzeniami zagregowana metoda dopasowania rozkładu może wskazać inny rozkład jako najlepiej dopasowany gdyż nowe dane, zwłaszcza gdy będzie ich dużo w stosunku do dotychczas uwzględnionych, mogą układać się wg innego modelu. Bardzo pomocna w takim przypadku jest analiza siatki rozkładu prawdopodobieństwa, pozwalająca wstępnie ocenić dopasowanie wybranego modelu hipotetycznego do odpowiedniego przypadku.

Analizując dane o uszkodzeniach z użyciem zagregowanej metody, należy zdawać sobie sprawę z tego, że czasami żaden z modeli statystycznych rozkładu nie pasuje do analizowanych danych. W takim przypadku otrzymuje się najlepsze rozwiązanie z najgorszych, które może słabo opisywać dane (niewielkie dopasowanie). W innych przypadkach, w których wiele modeli może być dobrze dopasowanych do danych empirycznych, same statystyki nie wystarczą, w takim przypadku znajomość mechanizmu uszkodzenia może być nieoceniona przy doborze najbardziej odpowiedniego modelu hipotetycznego. Ważne jest, aby pamiętać, że chociaż zagregowana metoda dla małych prób, także uszereguje wybrane rozkłady prawdopodobieństwa w zależności od liczby parametrów w konkretnym rozkładzie hipotetycznym, to używanie jej w takich przypadkach, obarczone jest dużą niepewnością i zalecane jest aby używać jej dla większych zbiorów danych.

Należy ponadto mieć na uwadze, że dwuparametrowy rozkład wykładniczy, trójparametrowy Weibulla oraz uogólniony gamma, zawierają parametr położenia, którego zmiana powoduje przesunięcie dystrybuanty i funkcji rozkładu prawdopodobieństwa bez zmiany jego kształtu.

Uogólniony rozkład gamma to natomiast złożony model, który może łatwo naśladować wiele innych rozkładów i dlatego często wydaje się być najlepiej dopasowanym do analizowanych danych.

Analizując dane naniesione na siatki probabilistyczne, często można stwierdzić, że mają one więcej niż jeden rodzaj uszkodzenia (np.: zmęczeniowe, eksploatacyjne, konstrukcyjne, technologiczne, itd.). W takim przypadku wszystkie rozkłady uszeregowane według zagregowanej metody wyboru rozkładu mogą okazać się źle dopasowanymi, ponieważ opracowana metoda może być stosowana tylko dla jednorodnego typu uszkodzeń badanego elementu. W takich sytuacjach należy uwzględnić możliwość skorzystania z mieszaniny jednego lub kilku rozkładów, np. z połączenia dwóch rozkładów Weibulla.

6. Podsumowanie

Uzyskane wyniki stanowią ważny argument do możliwości zastosowania proponowanej zagregowanej metody doboru rozkładu hipotetycznego do danych empirycznych. Uwzględniając trzy kryteria oceny dokładności dopasowania, unika się błędów jakie można popełnić uzależniając się tylko od jednego z nich.

Używanie tylko jednego kryterium określającego jakość dopasowania dystrybuanty hipotetycznej do empirycznej, często może okazać się niewystarczające, zależy to od wielu zmiennych głównie od ilości danych, czy dane są pełne czy ucięte (cenzurowane), ale przede wszystkim od rodzaju uszkodzenia.

Zagregowana metoda identyfikacji rozkładu hipotetycznego uwzględniająca trzy kryteria, jest metodą ogólną i ma szerokie zastosowanie, przy spełnieniu odpowiednich warunków, a więc liczba obserwacji musi być odpowiednio duża oraz powinna zawierać dokładne dane dotyczące czasów do uszkodzenia lub do zakończenia badań. Zmodyfikowana statystyka K-S (AVGOV) jest wrażliwa na lokalne odchylenia. Z kolei średnie odchylenie bezwzględne dystrybuanty hipotetycznej od dystrybuanty empirycznej (AVPLOT), nie jest już tak wrażliwe na lokalne odchylenia, uwzględnia natomiast globalne zróżnicowanie rozkładów i stanowi dobre uzupełnienie do kryterium MK-S. Dla trzeciego kryterium, czyli logarytmu funkcji wiarygodności (LKV) istotna jest wielkość próby, gdyż w przypadku małych próbek uzyskana wartość może być mocno obciążona.

Korzyści wynikające z poprawnego doboru rozkładu zmiennej losowej czasu zdatności obiektu odnawialnego jakim jest pojazd szynowy są znaczące m.in. ze względu na koszty generowane poprzez niewykorzystanie trwałości potencjalnej elementu jak i strat wynikających z wymuszonych obsług korekcyjnych i nieplanowanego przestoju pojazdu.

Badania zostały sfinansowano z następujących środków: PUT 04/43/SBAD/0111 oraz 05/51/SBAD/3584

Literatura

1. Abernethy R B. The New Weibull Handbook: Reliability & Statistical Analysis for Predicting Life, Safety, Survivability, Risk, Cost, and Warranty Claims (Fifth ed.), Florida, 2010.
2. Andrzejczak K, Selech J. Flexible Prediction of the Vehicle Component Damage. Transport Means 2018: Proceedings of the 22nd International Scientific Conference, Trakai, Lithuania, Part II, 2018; 987-990.
3. Andrzejczak K, Selech J. Generalised Gamma Distribution in the Corrective Maintenance Prediction of Homogeneous Vehicles. In: Kabashkin I, Yatskiv (Jackiva) I, Prentkovskis O. (eds) Reliability and Statistics in Transportation and Communication. RelStat 2018. Lecture Notes in Networks and Systems. Springer, Cham 2018; 68: 519-530, ISBN 978-3-030-12450-2.

4. Andrzejczak K, Selech J. Investigating the trends of average costs of corrective maintenance of public transport vehicles. Journal of KONBiN, 2017; 41: 207-226. DOI 10.1515jok-2017-0011.
5. Andrzejczak K, Selech J. Quantile analysis of the operating costs of the public transport fleet. Transport Problems, 2017; 12 (3): 103-111.
6. Andrzejczak K. Statystyka elementarna z wykorzystaniem systemu Statgraphics, Wyd. Politechniki Poznańskiej, Poznań 1997.
7. Bartnik G, Pieniak D, Niewczas A M, Marciniak A. Probabilistic model for flexural strength of dental composites used in modelling reliability of the "tooth-dental composite" system. Eksploatacja i Niezawodnosc – Maintenance and Reliability, 2016; 18 (1): 136–141, http://dx.doi.org/10.17531/ein.2016.1.18.
8. Bavuso S J. Aerospace Applications of Weibull and Monte Carlo Simulation with Importance Sampling, IEEE, Annual Reliability and Maintainability Symposium, Proc. 1997.
9. Dolce J E. Analytical Fleet Maintenance Management, SAE International, SUA, 1994.
10. Elmahdy E E. Modelling Reliability Data with Finite Weibull or Lognormal Mixture Distributions. Appl. Math. Inf. Sci. 2017; 11 (9), 1081-1089.
11. Ferreira L A, Silva J L. Parameter estimation for Weibull distribution with right censored data using EM algorithm. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (2): 310–315, http://dx.doi.org/10.17531/ein.2017.2.20.
12. Fuc P, Rymaniak L, Ziolkowski A. The correlation of distribution of PM number emitted under actual conditions of operation by PC and HDV vehicles, WIT Transactions on Ecology and the Environment. WIT Press, 2013; 174: 207. ISBN: 978-1-84564-718-6.
13. Gill A. Optimisation of the technical object maintenance system taking account of risk analysis results. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 420–431, http://dx.doi.org/10.17531/ein.2017.3.13.
14. Hajkowski J, Popielarski P, Sika R. Prediction of HPDC casting properties made of AlSi9Cu3 alloy, Advances in Manufacturing, SPRINGER, Manufacturing, 2017, 621-631, DOI: 10.1007/978-3-319-68619-6_59.
15. Hirose H. Bias Correction for the Maximum Likelihood Estimation in Two-parameter Weibull Distribution, IEEE Transactions on Dielectrics and Electrical Insulation, 1999; 6, 1,
16. https://www.reliasoft.com/Weibull [dostęp 2018].
17. Johnson R A, Miller I, Freund J E. Probability and Statistics for Engineers, eighth ed., Pearson Education Limited Co., UK, 2014.
18. Kececioglu D. Reliability & Life Testing Handbook, PrenticeHall, Inc., Englewood Cliffs, New Jersey, 1993; 1.
19. Lawless J F. Statistical Models and Methods for Lifetime Data, second ed., Wiley, 2002.
20. Lawless J F. Statistical Models And Methods for Lifetime Data, John Wiley & Sons, Inc., New York, 1982.
21. Lee E T, Wang J W. Statistical Methods for Survival. Data Analysis, John Wiley & Sons Inc; (3rd Edition), 2003.
22. Legát V, Mošna F, Aleš Z, Jurča V. Preventive maintenance models – higher operational reliability. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (1): 134–141, http://dx.doi.org/10.17531/ein.2017.1.19.
23. Liu J, Song B, and Zhang Y. Competing failure model for mechanical system with multiple functional failures. Advances in Mechanical Engineering 2018, 10(5) 1–16. DOI: 10.1177/1687814018773155.
24. Loska A. Exploitation assessment of selected technical objects using taxonomic methods, Eksploatacja i Niezawodność − Maintenance and Reliability, 2013; 15, 1.
25. Młynarski S, Pilch R, Smolnik M, Szybka J. Methodology of network systems reliability assessment on the example of urban transport. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 278–283, http://dx.doi.org/10.17531/ein.2018.2.14.

26. Młyńczak M. Analiza danych eksploatacyjnych w badaniach niezawodności obiektów technicznych, Zeszyty Naukowe WSOWL, 2001; 1 (159).
27. Nelson W. Applied Life Data Analysis, John Wiley & Sons, Inc., New York, 1982.
28. Perz P, Malujda I, Wilczyński D, Tarkowski P. Methods of controlling a hybrid positioning system using LabVIEW, 21th Scientific Polish–Slovak Conference ”Machine Modeling and Simulations 2016”, Procedia Engineering 2017; 177, 339-346.
29. Pieniak D, Niewczas A M, Niewczas A, Bieniaś J. Analysis of Survival Probability and Reliability of the Tooth-composite Filling System. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2011; 2(50): 25-34.
30. ReliaSoft Corporation, Life Data (Weibull) Analysis Reference, ReliaSoft Publishing Tucson, AZ, 2008.
31. Research Project “Increase in the efficiency of functioning of public means of transport as a result of implementation of LCC and RAMS concepts in accordance with the IRIS standards based on integrated information technology system” financed by Polish National Center for Research and Development. No. PBS3/B6/30/2015.
32. Rojek I, Kujawińska A, Hamrol A, Rogalewicz M. Artificial neural networks as a means for making process control charts user friendly. In: Burduk A., Mazurkiewicz D. (eds.), Intelligent Systems in Production Engineering and Maintenance – ISPEM 2017, Advances in Intelligent Systems and Computing, Springer, 637, 168-178, 2017.
33. Selech J. Prognozowanie kosztów obsługiwania korekcyjnego pojazdów transportu masowego. Monografia, Wydawnictwo Naukowe ITeE-PIB, Radom 2019, ISBN 978-83-7789-557-3.
34. Świderski A, Jóźwiak A, Jachimowski R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks, Eksploatacja i Niezawodność – Maintenance and Reliability, 2018; 20 (2), 292-299.
35. Trojanowska J, Kolinski A, Galusik D, Varela M L R, Machado J. A methodology of improvement of manufacturing productivity through increasing operational efficiency of the production process. In: Hamrol A., Ciszak O., Legutko S., Jurczyk M. (eds) Advances in Manufacturing. Lecture Notes in Mechanical Engineering. Springer, Cham, 2018; 23-32.
36. Walus K J. Driver's Strategy and Braking Distance in Winter, Transport Means 2017: Proceedings of the 21st International Scientific Conference, Juodkrante, Lithuania. 2017; Part 2, 505 – 509, ISSN 1822-296 X, e-ISSN 2351-7034.
37. Wojtkowiak D, Talaśka K, Malujda I, Domek G. Estimation of the perforation force for polymer composite conveyor belts taking into consideration the shape of the piercing punch. The International Journal of Advanced Manufacturing Technology. 2018. https://doi.org/10.1007/s00170-018-2381-3.
38. Ziółkowski J, Borucka A. Model Markowa w logistycznym zarządzaniu przedsiębiorstwem, Journal of Conbin, 2016; 2 (38). ISSN 1895-82-82
39. Żurek J, Ziółkowski J, Borucka A. Application of Markov processes to the method for analysis of combat vehicle operation in the aspect of their availability and readiness, Safety and Reliability – Theory and Applications – Čepin & Briš (Eds)©, Taylor & Francis Group, London, 2017; 2343-2352. ISBN 978-1-138-62937-0.

kpt. dr inż. Arkadiusz Jóźwiak

Wydział Bezpieczeństwa, Logistyki i Zarządzania
Wojskowa Akademia Techniczna
ul. gen. Sylwestra Kaliskiego 2, 00-908 Warszawa 46, Polska
e-mail: arkadiusz.jozwiak@wat.edu.pl

mgr inż. Paulina Owczarek

Wydział Bezpieczeństwa, Logistyki i Zarządzania
Wojskowa Akademia Techniczna
ul. gen. Sylwestra Kaliskiego 2, 00-908 Warszawa 46, Polska
e-mail: paulina.owczarek@wat.edu.pl

prof. dr hab. inż. Leon Prochowski

Wydział Inżynierii Mechanicznej
Wojskowa Akademia Techniczna
ul. gen. Sylwestra Kaliskiego 2, 00-908 Warszawa 46, Polska
e-mail: leon.prochowski@wat.edu.pl

dr hab. inż. Andrzej Świderski, prof. ITS

Instytut Transportu Samochodowego
ul. Jagiellońska 80, 03-301 Warszawa
e-mail: andrzej.swiderski@its.waw.pl

Badanie wpływu czasu wykorzystania samochodów kategorii N1 na efektywność ekonomiczną ich eksploatacji

Słowa kluczowe: *eksploatacja samochodów, pojazdy samochodowe kategorii N1, efektywność ekonomiczna, sieci neuronowe*

Streszczenie: Rozważa się efektywność eksploatacji samochodów ciężarowych o DMC < 3,5 tony. Są to pojazdy należące do kategorii N1 (według Dyrektywy 2007/46/WE) zwykle nazywane samochodami dostawczymi. Do prowadzonej analizy efektywności wykorzystano wyniki obserwacji z realizacji zleceń przewozowych w 7 firmach transportowych z sektora MŚP. Grupa badawcza objęła 24 pojazdy, które wykonywały zadania transportowe w strefie miejskiej i w najbliższym otoczeniu miasta. Informacje gromadzono w cyklach miesięcznych.

Podczas analizy efektywności ekonomicznej zastosowano kilka miar przychodu (bezwzględny i względny). Obliczenia prowadzono przy wykorzystaniu modelu procesu eksploatacji pojazdów w postaci sieci neuronowej, w której brano pod uwagę zbiór 12 zmiennych wejściowych i 3 zmienne wyjściowe. Stosując program komputerowy Statistica 13.3 oraz zdefiniowanie grupy i czynniki opisujące proces realizacji poszczególnych zadań transportowych, opracowany model sieci neuronowej umożliwił poszukiwanie wpływu wybranych czynników eksploatacyjnych na efektywność ekonomiczną samochodów kategorii N1.

Przeprowadzone obliczenia pokazały istotny wpływ liczby dni pracy pojazdów w miesiącu, masę ładunku, a także porę roku. Uzyskane wyniki obliczeń pokazały specyficzne cechy wpływu liczby dni pracy na przychód w firmie transportowej. Wzrost liczby dni pracy sprzyja wzrostowi przychodu w sposób ograniczony, a to ograniczenie zależy m.in. od pory roku.

1. Wstęp

Specyfika użytkowania pojazdów samochodowych kategorii N1 stale wzbudza wiele kontrowersji, a w obszarze uwarunkowań prawnych nadal pojawia się szereg niejasności, co wpływa na znikomą literaturę w tym obszarze. W każdym miesiącu pojawiają się nowe informacje dotyczące statystyk Głównego Inspektoratu Transportu Drogowego o wynikach kontroli pojazdów samochodowych kategorii N1. Od kilku lat udział procentowy pojazdów z ładunkiem przekraczającym ich dopuszczalną ładowność w stosunku do skontrolowanych pojazdów kategorii N1 utrzymuje się na poziomie średnio 93% (tabela 1) [19]. Faktem jest, że Inspekcja Transportu Drogowego kontroluje zwykle pojazdy ciężarowe powyżej 3,5 tony, natomiast pojazdy kategorii N1 tylko wtedy, kiedy uzna, że istnieje wyraźne podejrzenie popełnienia konkretnego wykroczenia. Liczba kontrolowanych pojazdów jest znikoma, nie mniej jednak wskaźnik udziału procentowego pojazdów z ładunkiem przekraczającym ich dopuszczalną ładowność do 3,5 tony wskazywać może na istnienie złożonego problemu, którego poddać należy szczegółowej analizie.

Tabela 1. Udział procentowy pojazdów z ładunkiem przekraczającym ich dopuszczalną ładowność w stosunku do skontrolowanych pojazdów kategorii N1 [19]

	2014 r.	2015 r.	2016 r.	2017 r.
Liczba wszystkich pojazdów o DMC do 3,5 t	2 399 323	2 447 764	2 515 751	2 574 312
Liczba skontrolowanych pojazdów o DMC do 3,5 t	6 685	6 507	7 594	10 048
Liczba wystawionych mandatów pojazdom z ładunkiem przekraczającym ich dopuszczalną ładowność	6 135	6 172	6 172	9 396
Udział procentowy	92%	95%	92%	94%

Zasadne staje się więc zainteresowanie tematyką zwiększania zysków z eksploatacji pojazdów samochodowych kategorii N1. Odwzorowanie i uproszczenie zjawisk rzeczywistych w postaci modelu staje się ważnym elementem w poszukiwaniu skutecznych metod opisu problemów i zakłóceń w procesie użytkowania i obsługiwania pojazdów samochodowych. W rezultacie ułatwia to znalezienie sposobów na zwiększanie efektywności ekonomicznej przedsiębiorstw transportowych.

W artykule podmiotem badań jest efektywność eksploatacji samochodów kategorii N1, które zgodnie z [27] określa się jako pojazdy zaprojektowane i wykonane do przewozu ładunków i mające maksymalną masę całkowitą nieprzekraczającą 3,5 tony (DMC do 3,5 t). W praktyce tego typu pojazdy określa się pojęciem pojazdy dostawcze, dlatego w dalszej części artykułu pojęcie to stosowane będzie zamiennie.

Planowanie i realizacja procesu eksploatacji pojazdów samochodowych w złożonym systemie transportowym związane jest z rozwiązywaniem wielokryterialnych problemów decyzyjnych, które skupiają się m.in. na minimalizacji kosztów i osiągnięciu maksymalnego zysku. Zagadnienie to dotyczy problematyki dwóch podstawowych elementów procesu

eksploatacji, tj. użytkowania i obsługiwania pojazdów samochodowych. Efektywne wykorzystanie środków transportu w każdym przedsiębiorstwie jest jednym z głównych sposobów na osiągnięcie przewagi konkurencyjnej.

Szeroko prowadzone analizy procesu eksploatacji pojazdów samochodowych dotyczą najczęściej minimalizowania kosztów związanych z ich użytkowaniem oraz zapewniania maksymalnej niezawodności systemu transportowego, a także wpływu użytkowania pojazdów na środowisko naturalne [21], czy też aspektu bezpieczeństwa [20]. Natomiast ocena intensywności eksploatacji pojazdów samochodowych jest prowadzona z uwzględnieniem m.in. wartości przebiegu, pojemności silnika, wieku pojazdu [18], kosztów napraw, przychodów [16], gotowości technicznej, intensywności użytkowania pojazdów samochodowych [6]. Ze względu na losowy charakter awarii pojazdów dla utrzymania efektywnej i bezpiecznej ich eksploatacji niezbędna okazuje się wiedza dotycząca procesów stochastycznych [2].

Modelowanie i ocena tak złożonych procesów na podstawie klasycznych modeli matematycznych oraz techniki teorii niezawodności może być trudna do przeprowadzenia i nie przynosić oczekiwanych rezultatów ze względu na dużą liczbę danych ilościowych i jakościowych oraz ze względu na dynamicznie zmieniające się uwarunkowania systemu eksploatacji pojazdów. W takim przypadku proponowane są inne metody obliczeniowe, np. modele wykorzystujące procesy Markowa lub niezawodnościowe diagramy fazowe, model sieci Petriego lub procesy symulacji Monte Carlo [13], algorytm odpornościowej selekcji klonalnej [5].

Biorąc pod uwagę stopień złożoności badanego problemu oraz coraz szersze wykorzystanie metod sztucznej inteligencji do rozwiązywania tego typu zadań, celem niniejszej pracy staje się ocena efektywności ekonomicznej przedsiębiorstw transportowych przy wykorzystaniu sieci neuronowej. Ocena jest ukierunkowana na firmy transportowe, funkcjonujące na rzecz transportu drobnicowego w obszarach miejskich i podmiejskich. W pracy wykorzystano kilka miar przychodu w działalności transportowej opisane w dalszej części pracy.

Wszędzie tam, gdzie nie ma podstaw do aproksymacji liniowej występujących zjawisk i procesów, zwykle przy rozwiązywaniu trudnych i kłopotliwych zagadnień oceny m.in. efektywności eksploatacji samochodów, racjonalnym może być odwołanie się do sieci neuronowych lub innych algorytmów sztucznej inteligencji (a więc modeli, które odwzorowują zależności nieliniowe) [7], [9], [10], [24]. Sztuczne sieci neuronowe są jedną z technik wykorzystywaną przez sztuczną inteligencję. Istnieją również inne zastosowania sztucznej inteligencji w transporcie. Dla przykładu: do oceny zapewnienia jakości środków transportu, do optymalizacji tras przejazdu [11], czy do oceny zarządzania konfiguracją usług transportowych [23].

2. Analiza kosztów przedsiębiorstwa transportowego z taborem pojazdów samochodowych kategorii N1

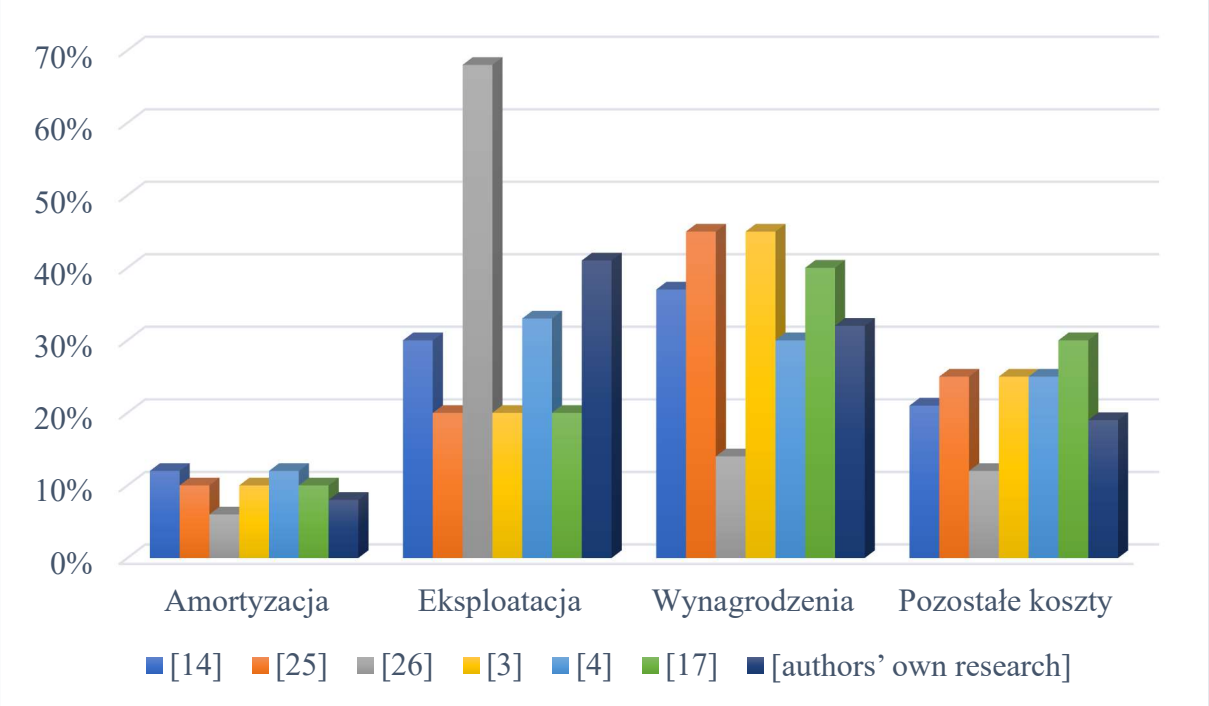
W dalszym ciągu w pierwszej kolejności wyboru dostawcy usług transportowych wybiera się poprzez kryterium minimalnej ceny. Duża konkurencja i stale rosnące wymagania rynku transportowego zmuszają przewoźników do nieustannego poszukiwania metod minimalizowania kosztów przedsiębiorstwa przewozowego. Generowanie przychodów w firmach transportowych opiera się przede wszystkim na intensywności użytkowania pojazdów. Z reguły są one proporcjonalne do liczby przejechanych kilometrów, masy ładunku czy przepracowanych godzin. Intensywne użytkowanie pojazdów generuje nie tylko wzrost przychów, ale też kosztów, dlatego niezmiernie ważne jest dokonanie ich szczegółowej analizy.

W przedsiębiorstwach, także transportowych, jedną z najczęściej stosowanych metod podziału kosztów jest ich układ rodzajowy, zawierający 7 grup, które stanowią jednocześnie nazwy kont syntetycznych: amortyzacja, zużycie materiałów i energii, usługi obce, podatki i opłaty, wynagrodzenia, ubezpieczenia społeczne i inne świadczenia, pozostałe koszty (rodzajowe).

Wg wielu publikacji koszty usług obcych stanowią największy udział procentowy w stosunku do wszystkich kosztów przedsiębiorstwa transportowego [15]. W niniejszym opracowaniu podmiotem badań są *mikro, małe i* średnie przedsiębiorstwa (mikro i MŚP), dlatego struktura kosztów będzie nieco różniła się od ogólnej klasyfikacji kosztów rodzajowych przedsiębiorstw. Powodem tego może być chociażby fakt posiadania przez mikro i MŚP jedynie własnego, nieleasingowanego już taboru, który usługi przewozowe świadczy bezpośrednio bez udziału firm outsourcingowych. Zagadnienia związane szczegółowo z kosztami przedsiębiorstw transportu samochodowego ładunków są przedmiotem zainteresowań wielu autorów [1], [4], [8], [12], [15], [26] którzy najczęściej sprowadzają je do czterech podstawowych grup rodzajowych i określają ich wartości procentowe w stosunku do pozostałych kosztów:

- amortyzacja 6% – 12%,
- eksploatacja 20% – 68%,
- wynagrodzenie kierowców 14% – 45 %,
- pozostałe koszty 12% – 30 %.

Na podstawie przytoczonej analizy literatury, na rys. 1 przedstawiono udziały trzech podstawowych grup kosztów przedsiębiorstw samochodowego transportu ładunków z perspektywy kilku autorów.



Rys. 1. Porównanie udziału wybranych grup kosztów przedsiębiorstw samochodowego transportu ładunków

Na potrzeby osiągnięcia celu badawczego zebrano dane, które sklasyfikowano w czterech grupach czynników: użytkowe, pory roku, obsługowe, ekonomiczne.

3. Metoda i obiekt badań

Badaniom poddano zadania wykonywane w firmach transportowych, które realizują usługi na terenie Polski. Praca taboru, należącego do 7 różnych firm transportowych z sektora MŚP polega na realizacji zadań transportowych zgodnie z zapotrzebowaniem klientów. Grupa badawcza obejmuje 24 pojazdy samochodowe kategorii N1, 5 modeli: Renault Master, Renault Mascott, Citroen Jumper i Fiat Ducato. W badaniach brano pod uwagę jedynie te dane techniczne, które miały wpływ na wymienione wcześniej czynniki.

Zdefiniowano następujące miary efektywności ekonomicznej:

- przychód $[Y_M^E]$ – wyrażony jako różnica miesięcznej wartości i miesięcznych kosztów eksploatacji,
- przychód względny Y_L^E – wyrażony jako stosunek miesięcznej wartości zleceń do przebiegu miesięcznego,
- zysk względny Y_W^E – wyrażony jako stosunek dochodu do przebiegu miesięcznego.

Zdefiniowano 4 główne grupy czynników branych pod uwagę w badaniach:

- Y^U - grupa wielkości opisujących czynniki użytkowania pojazdów samochodowych: charakteryzujące sposób i intensywność wykonywanej pracy,
- Y^K - grupa wielkości opisujących pory roku: definiujące warunki zewnętrzne w jakich użytkowany jest pojazd,
- Y^O - grupa wielkości opisujących czynności obsługowe pojazdów samochodowych: dotyczące prowadzonej strategii obsługowej oraz jej efektów,
- Y^E - grupa wielkości opisujących czynniki ekonomiczne: związane z kosztami i rentownością realizacji zadań przewozowych.

Zbiór czynników opisujących ww. grupy przedstawiono w tabela 2.

Tabela 2. Zbiór czynników wykorzystanych podczas modelowania procesu eksploatacji

	Oznaczenia grup i czynników		**Jednostki miary**
	$\boldsymbol{Y^U}$	**Grupa: czynniki użytkowania pojazdów samochodowych**	
1	Y_D^U	liczba dni użytkowania pojazdu w miesiącu	liczba
2	Y_R^U	miesięczny przebieg pojazdu	kilometry
3	Y_J^U	miesięczny czas jazdy pojazdu	minuty
4	Y_C^U	miesięczny czas pracy pojazdu	minuty
5	Y_P^U	średnie zużycie paliwa	litry/100 km
6	Y_M^U	średnia dzienna masa ładunku	kilogramy
7	Y_E^U	wartość procentowa wykorzystania ładowności	%
	$\boldsymbol{Y^K}$	**Grupa: pora roku**	
8	Y_W^K	pora roku	sezon 1, sezon 2, sezon 3
	$\boldsymbol{Y^O}$	**Grupa: czynności obsługowe pojazdów samochodowych**	
9	Y_P^O	uzupełnianie płynów	wykonano, nie wykonano
10	Y_K^O	serwis ogumienia	wykonano, nie wykonano
11	Y_H^O	serwis hamulców	wykonano, nie wykonano

	Oznaczenia grup i czynników		Jednostki miary
	Y^E	**Grupa: czynniki ekonomiczne**	
12	Y_Z^E	miesięczna wartość zleceń	zł
13	Y_K^E	miesięczny koszt eksploatacji	zł
14	Y_M^E	miesięczny przychód z realizacji usług przewozowych	zł
15	Y_L^E	względny jednostkowy przychód	zł/km
16	Y_W^E	względny zysk jednostkowy	zł/km

Dla tak zdefiniowanych grup zebrano rzeczywiste dane z jednego roku eksploatacji pojazdów (2017) w cyklach miesięcznych. Pojazdy te poruszały się głównie w ruchu miejskim z pojedynczymi trasami poza miastem na terenie kraju. Dane uzyskano ze zleceń transportowych w badanym okresie, analiz ekspertyz serwisowych i wywiadów z ekspertami (dyspozytorami, kierowcami, serwisantami, mechanikami). Dla każdego pojazdu dokonano 156 obserwacji wymienionych czynników. W taki sposób zebrano 3744 danych, które posłużyły do modelowania procesu eksploatacji z wykorzystaniem sieci neuronowej.

4. Modelowanie neuronowe

Podczas tworzenia sieci neuronowej wykorzystano część sygnałów z tabela 2, są to:

- wejściowe ilościowe: Y_D^U, Y_R^U, Y_J^U, Y_C^U, Y_P^U, Y_M^U, Y_E^U, Y_K^E,
- wejściowe jakościowe: Y_P^O, Y_K^O, Y_H^O, Y_W^K,
- wyjściowe ilościowe: Y_M^E, Y_L^E, Y_W^E.

Z wykorzystaniem wyników m.in. [22] pracy naukowej, zastosowano perceptron wielowarstwowy (*Multilayer Perceptron*) i algorytmy uczące: gradientów sprzężonych; najszybszego spadku i BFGS (*Broyden – Fletcher – Goldfarb - Shanno)*. W modelowaniu neuronowym przyjęto podział zbioru danych na części:

- 80 % - zbiór uczący wykorzystywany do modyfikacji wag,
- 10 % - zbiór testowy przeznaczony do bieżącego monitorowania procesu uczenia,
- 10 % - zbiór walidacyjny do oceny jakości sieci po zakończeniu procesu uczenia.

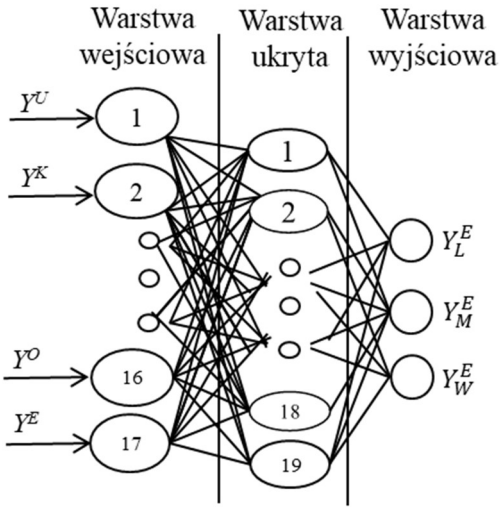
Po określeniu sygnałów wejściowych, sygnałów wyjściowych i parametrów sieci, przeprowadzono proces uczenia sieci neuronowej z wykorzystaniem programu komputerowego Statistica 13.3. Przykładowe jego wyniki przedstawiono w tabela 3.

Tabela 3. Przykładowe wyniki procesu uczenia sieci neuronowej

ID	Nazwa sieci	Jakość uczenia	Jakość testowania	Jakość walidacji	Algorytm uczenia	Aktywacja ukryta	Aktywacja wyjściowa
1	MLP 17-11-3	0,796875	0,746498	0,874968	BFGS 28	Wykładnicza	Wykładnicza
2	MLP 17-16-3	0,785931	0,764052	0,874283	BFGS 50	Sinus	Logistyczna
3	MLP 17-36-3	0,775433	0,804533	0,828288	BFGS 17	Sinus	Wykładnicza
4	MLP 17-30-3	0,762021	0,760954	0,861609	BFGS 23	Sinus	Liniowa
5	MLP 17-9-3	0,649014	0,612412	0,770544	BFGS 12	Sinus	Tanh
6	MLP 17-6-3	0,755626	0,766049	0,867752	BFGS 51	Liniowa	Tanh
7	MLP 17-8-3	0,767523	0,742646	0,876113	BFGS 66	Sinus	Tanh
8	MLP 17-31-3	0,776866	0,784677	0,872670	BFGS 36	Liniowa	Logistyczna
9	MLP 17-26-3	0,730811	0,783034	0,825576	BFGS 10	Wykładnicza	Sinus
10	MLP 17-9-3	0,807437	0,754784	0,865342	BFGS 38	Logistyczna	Liniowa
11	MLP 17-26-3	0,794570	0,754735	0,874439	BFGS 24	Tanh	Logistyczna
12	MLP 17-3-3	0,796772	0,816948	0,850614	BFGS 52	Logistyczna	Sinus
13	MLP 17-29-3	0,776284	0,783517	0,874145	BFGS 31	Liniowa	Logistyczna
14	MLP 17-19-3	0,813487	0,752890	0,844996	BFGS 45	Logistyczna	Sinus
15	MLP 17-23-3	0,776775	0,783067	0,873542	BFGS 32	Liniowa	Logistyczna
16	MLP 17-17-3	0,838238	0,761414	0,700263	BFGS 64	Tanh	Tanh
17	MLP 17-7-3	0,768890	0,799204	0,819687	BFGS 17	Liniowa	Wykładnicza
18	MLP 17-6-3	0,780286	0,768095	0,886923	BFGS 28	Logistyczna	Liniowa

5. Walidacja modelu sieci neuronowej i wyniki obliczeń

Struktura najlepszej sieci neuronowej przyjęła postać MLP 17-19-3, co oznacza 17 neuronów w warstwie wejściowej, 19 neuronów w warstwie ukrytej i 3 neurony w warstwie wyjściowej (rys. 2).



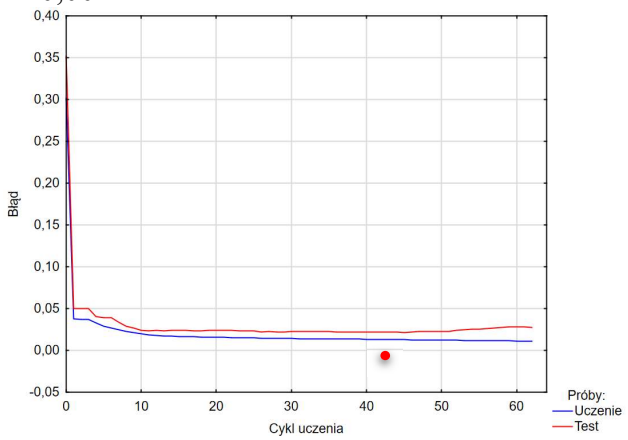
Rys. 2. Struktura utworzonej sieci MLP 17-19-3

Ponieważ wśród sygnałów wejściowych, pojawiły się sygnały wejściowe jakościowe, łączna liczba neuronów na wejściu stanowi sumę wszystkich sygnałów ilościowych i jakościowych z podziałem na poszczególne ich wartości. Tabela 4 przedstawia rozpisane sygnały wejściowe wybranej sieci neuronowej.

Tabela 4. Rozpisane sygnały wejściowe sieci neuronowej MLP 17-19-3

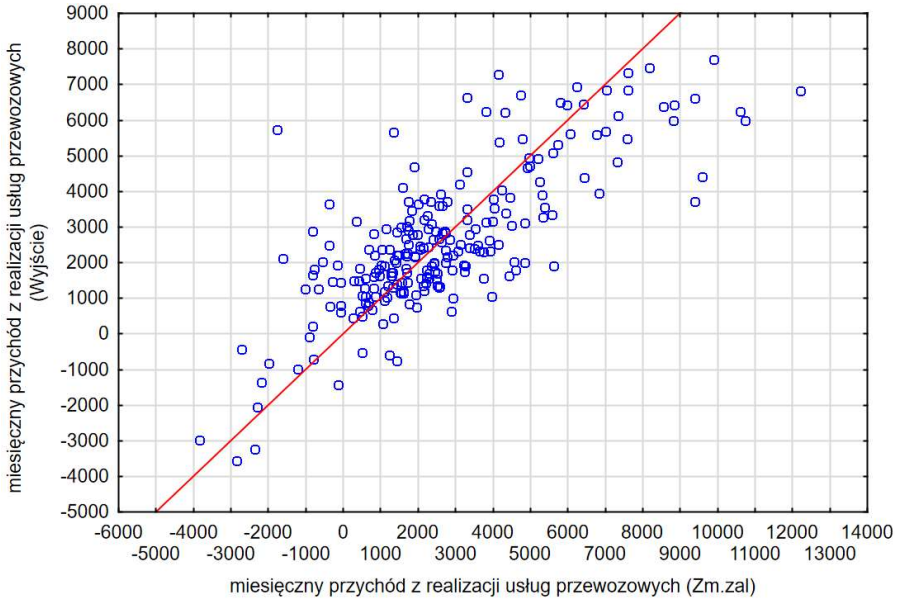
	Oznaczenia grup i czynników	Jednostki miary	ID neuronu	Wartość neuronu
1	Y_D^U	liczba	1	Y_D^U
2	Y_R^U	kilometry	2	Y_R^U
3	Y_J^U	minuty	3	Y_J^U
4	Y_C^U	minuty	4	Y_C^U
5	Y_P^U	litry/100 km	5	Y_P^U
6	Y_M^U	kilogramy	6	Y_M^U
7	Y_E^U	%	7	Y_E^U
8	Y_W^K	sezon 1, sezon 2, sezon 3	8	$Y_W^K\ sezon\ 1$
			9	$Y_W^K\ sezon\ 2$
			10	$Y_W^K\ sezon\ 3$
9	Y_P^O	wykonano, nie wykonano	11	$Y_P^O\ wykonano$
			12	$Y_P^O\ niewykonano$
10	Y_K^O	wykonano, nie wykonano	13	$Y_K^O\ wykonano$
			14	$Y_K^O\ niewykonano$
11	Y_H^O	wykonano, nie wykonano	15	$Y_H^O\ wykonano$
			16	$Y_H^O\ niewykonano$
12	Y_K^E	zł	17	Y_K^E

W tabela 3 jakość uczenia sieci MLP 17-19-3 została oszacowana na poziomie ok 81% prawdopodobieństwa wskazania poprawnej odpowiedzi, czyli przyjętej miary efektywności ekonomicznej. Jakość testowania - na poziomie 75% i jakość walidacji - na poziomie 85%. Najlepszym algorytmem uczenia okazał się algorytm BFGS 45. O pozytywnym wyniku uczenia sieci neuronowej świadczy m.in. wykres uczenia (rys. 3). Wynika z niego, że najlepszą strukturę sieci odnaleziono w 43 cyklu; udział błędnych odpowiedzi wynosił 19 %, a błąd został oszacowany na poziomie 0,002.



Rys. 3. Rezultaty uczenia sieci neuronowej MLP 17-19-3

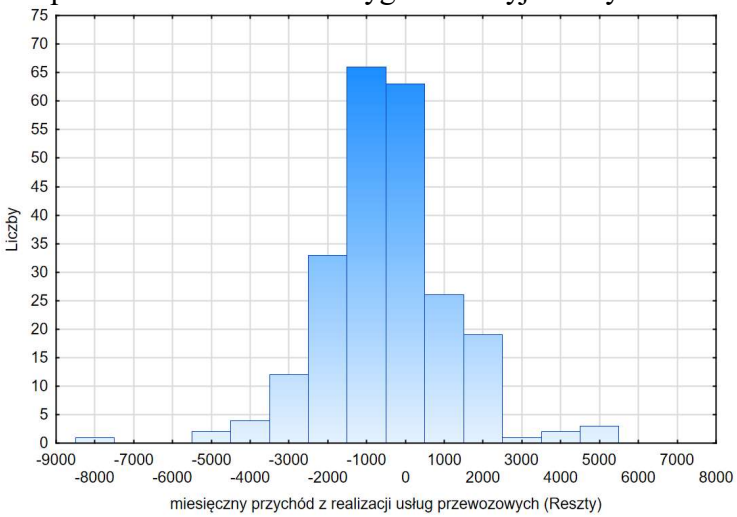
Również przebieg zmian rozrzutu, przestawiony na rys. 4, wskazuje na pozytywny wynik uczenia sieci.



Rys. 4. Rozrzut zmiennej zależnej sieci neuronowej MLP 17-19-3

Wyniki obliczeń przedstawionych na rys. 4 pokazują rozrzut między wartością prognozowaną przychodu (wynik obliczenia w sieci w trakcie uczenia) a jego rzeczywistą wartością. Widoczne skupienie wartości rozrzutu w pobliżu zera jest dobrym rezultatem obliczeń modelowych.

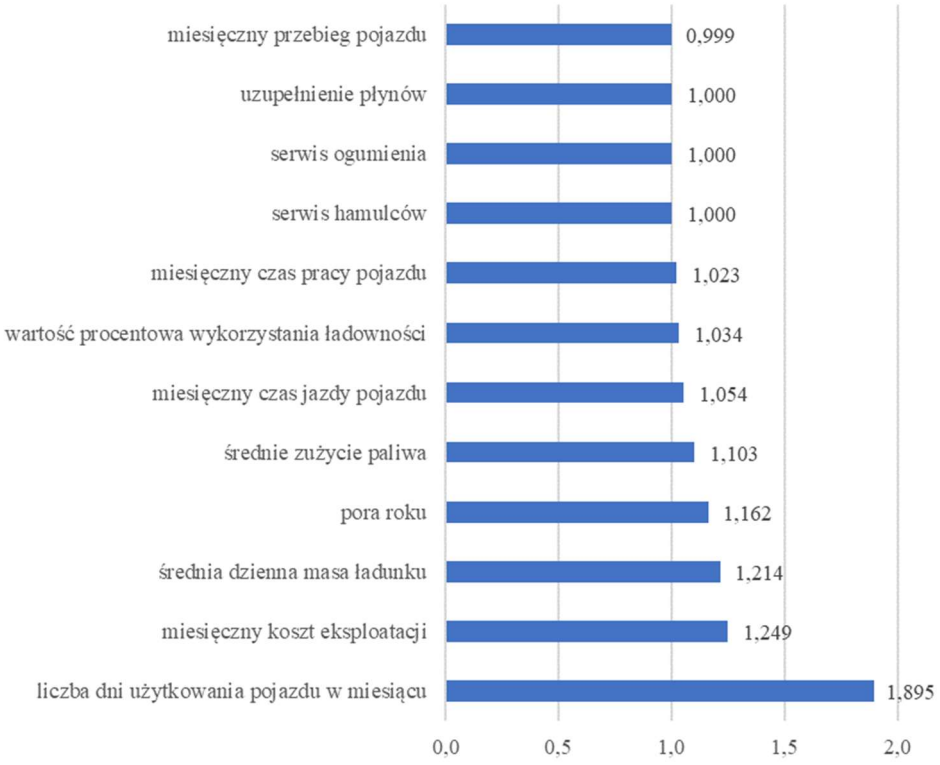
Histogram, przedstawiony na rys. 5 (rozkład reszt, czyli różnic między zmienną wyjściową i jej predykcją) pokazuje liczbę wyników obliczeń rozrzutu w pobliżu zera, co oznacza także wysoki poziom odwzorowania sygnałów wyjściowych.



Rys. 5. Rozkład reszt sieci neuronowej MLP 17-19-3

W kolejnym etapie, przeprowadzono analizę wrażliwości, która polegała na sprawdzeniu, jak zachowuje się błąd sieci w przypadku gdy modyfikowane są sygnały wejściowe. W tym obliczeniu wartości sygnału wejściowego zastępowane są przez średnią tego sygnału ze zbioru uczącego. Po podaniu tak zmodyfikowanych danych wejściowych sprawdzono błąd sieci. Jeśli błąd wzrósł znacznie, oznacza to, że sieć jest bardzo wrażliwa na dany sygnał.

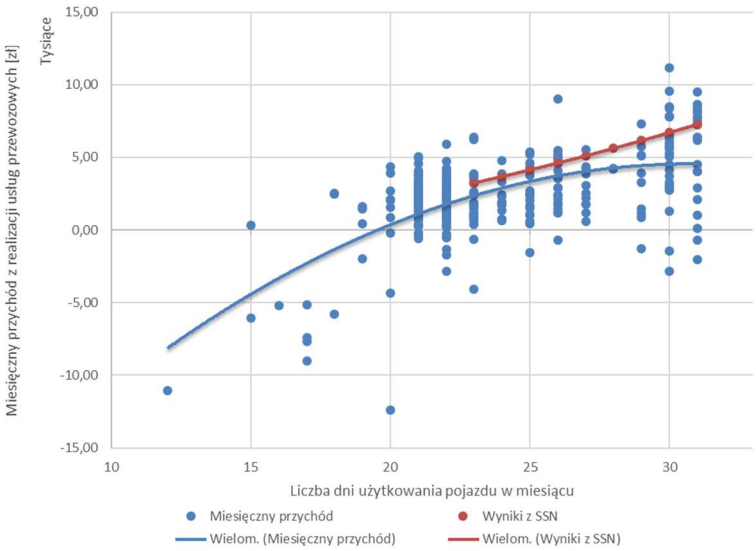
Globalna analiza wrażliwości odzwierciedla wpływ poszczególnych zmiennych wejściowych sieci na sygnały wyjściowe (rys. 6).



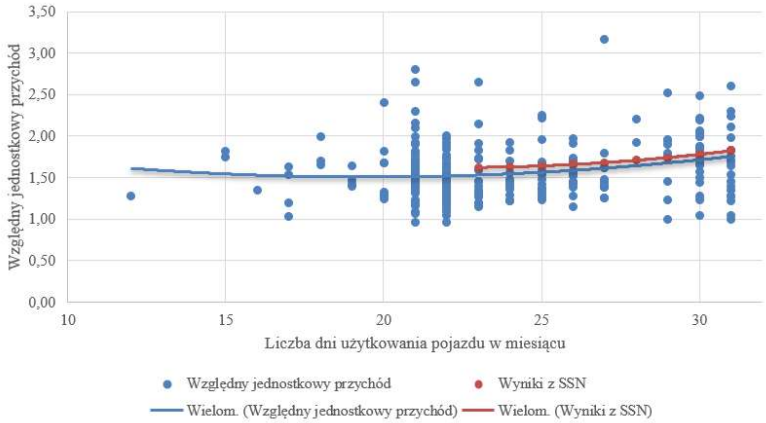
Rys. 6. Globalna analiza wrażliwości dla sieci neuronowej MLP 17-19-3

Obliczenia te pokazały, że największy wpływ na sygnały wyjściowe sieci neuronowej mają: miesięczna liczba dni pracy pojazdu, koszt eksploatacji, masa ładunku i sezon.

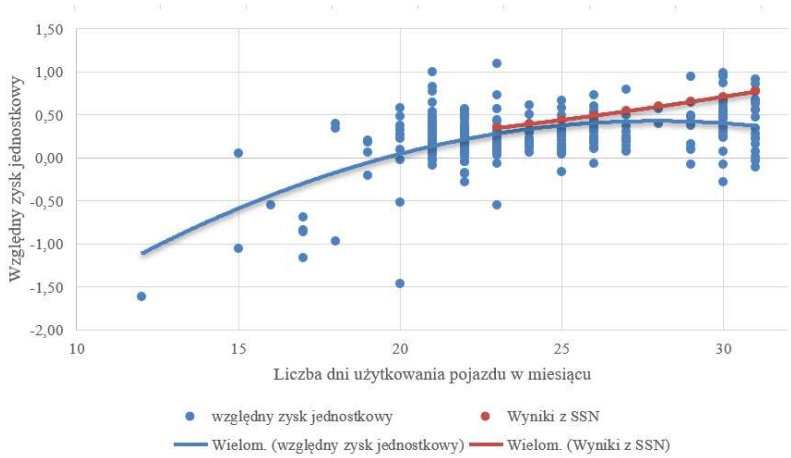
Na podstawie wybranej sieci neuronowej i danych zebranych do procesu uczenia sieci neuronowej pokazano trendy zmian wartości miar efektywności, mianowicie: miesięczny przychód, względny jednostkowy przychód i względny zysk jednostkowy w odniesieniu do liczby dni pracy pojazdu (rys. 7, rys. 8, rys. 9).



Rys. 7. Miesięczny przychód w odniesieniu do liczby dni pracy pojazdu



Rys. 8. Względny jednostkowy przychód w odniesieniu do liczby dni pracy pojazdu



Rys. 9. Względny zysk jednostkowy w odniesieniu do liczby dni pracy pojazdu

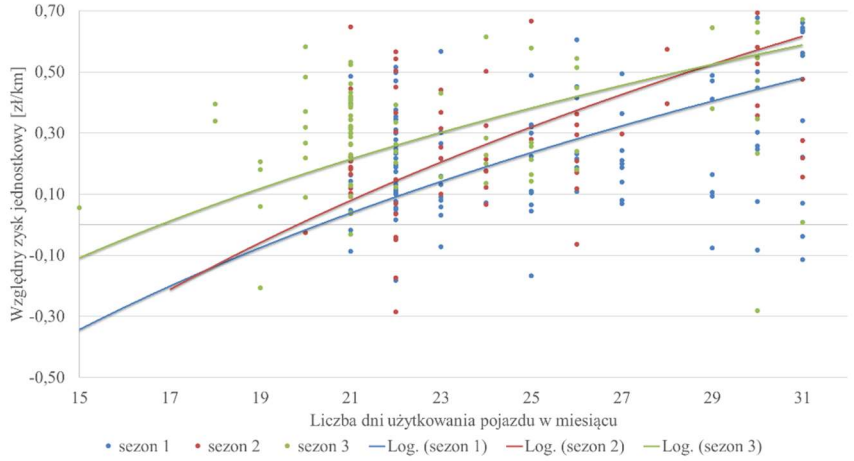
Z przeprowadzonej analizy również wynika, iż wybrana sieć prawidłowo odwzorowuje wybrane miary efektywności ekonomicznej.

W celu dokonania szczegółowej analizy wpływu tego czynnika na sygnał wyjściowy wyodrębniono wpływ pory roku na wartości miar efektywności w odniesieniu do liczby dni pracy pojazdu (rys. 10, rys. 11, rys. 12).

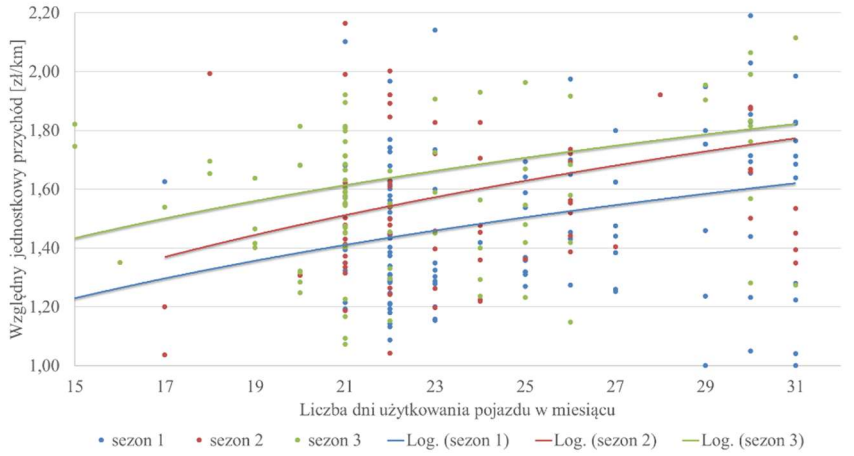
Na podstawie wywiadu z kierowcami pojazdów porę roku określono, jako ogólne warunki realizacji zleceń oraz komfort i bezpieczeństwo jazdy:

- sezon 1 przypisuje się miesiącom: maj, czerwiec, lipiec, sierpień,
- sezon 2: marzec, kwiecień, wrzesień, październik,
- sezon 3: styczeń, luty, listopad, grudzień.

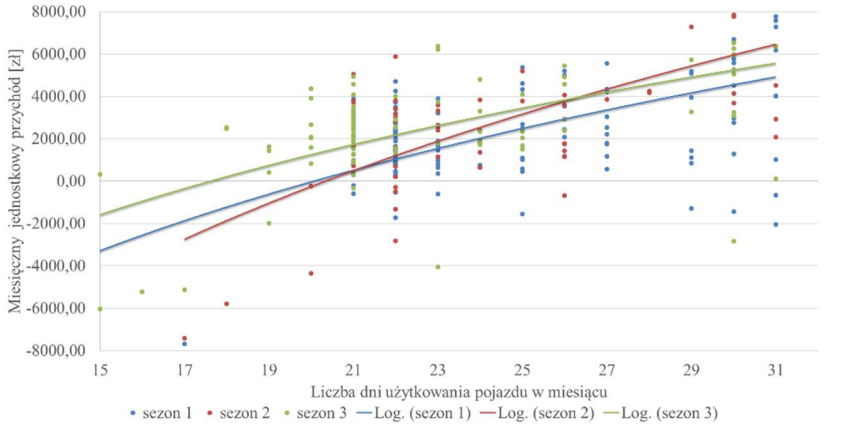
Przeprowadzone badania wykazały, że najwyższe wartości miar efektywności osiąga się podczas eksploatacji samochodów dostawczych i realizacji zleceń w czasie sezonu 3, a najniższe w czasie sezonu 1. Potwierdzeniem tego jest fakt występowania sezonowości w obszarze świadczenia usług transportowych. Złe warunki atmosferyczne, określające sezon zimowy, to czas wzmożonego użytkowania pojazdów ze względu na rosnące zapotrzebowanie na usługi (sezon 3), przy zmniejszonej ich ofercie. Dobre warunki atmosferyczne obserwuje się w sezonie letnim, ale wówczas maleje zapotrzebowanie na usługi przewozowe (sezon 1). Podaż usług w analizowanym sektorze w tym okresie jest wyższa od popytu na usługi.



Rys. 10. Względny zysk jednostkowy w odniesieniu do liczby dni pracy pojazdu w odniesieniu do Y^K



Rys. 11. Względny jednostkowy przychód w odniesieniu do liczby dni pracy pojazdu w odniesieniu do Y^K



Rys. 12. Miesięczny jednostkowy przychód w odniesieniu do liczby dni pracy pojazdu w odniesieniu do Y^K

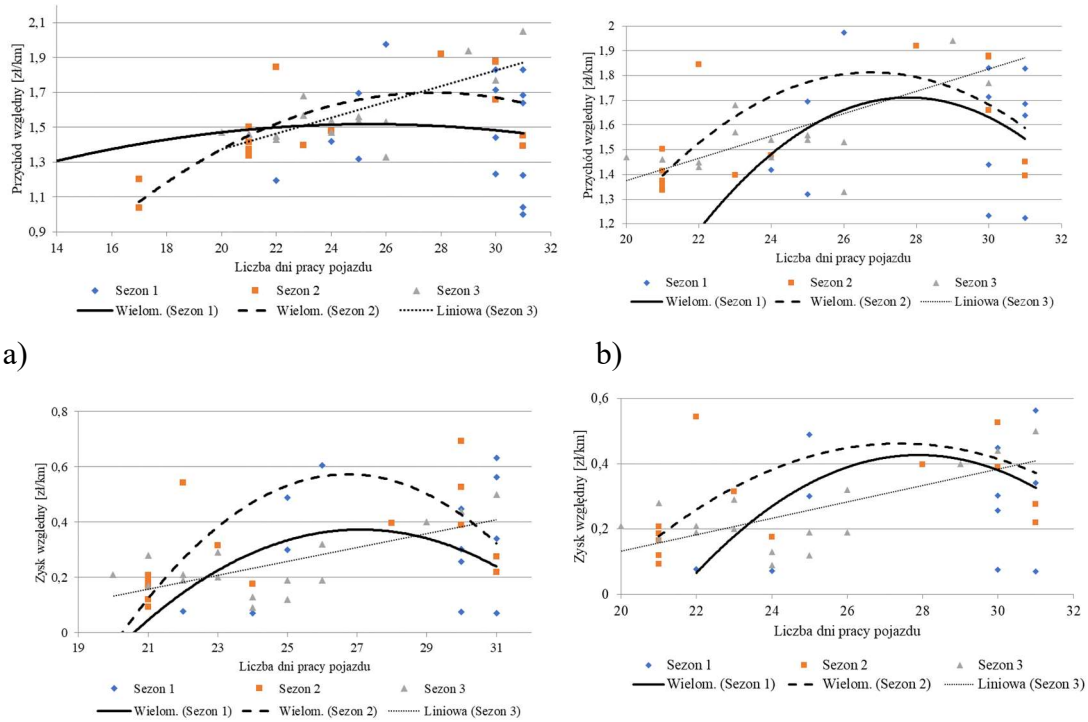
Weryfikację zaproponowanej metody przeprowadzono na podstawie wyników kolejnych obliczeń, dokonanych po wprowadzeniu do sieci neuronowej danych nie wykorzystanych w procesie uczenia. Przykładowe wyniki końcowe przedstawiono w tabela 5.

Tabela 5. Przykładowe wyniki obliczeń przychodu, jednostkowego przychodu względnego i jednostkowego zysku względnego dla zbioru danych nie wykorzystanych w procesie uczenia

P_Y^X	Y_D^U	Y_R^U	Y_J^U	Y_C^U	Y_P^U	Y_M^U	Y_E^U	Y_K^E	Y_W^K	Y_P^O	Y_K^O	Y_H^O	Y_M^E	Y_L^E	Y_W^E
1	30	11360	425	480	15	1661	1,33	12873,13	sezon 2	tak	tak	tak	5976,87	1,66	0,53
2	31	13331	516	553	15	1771	1,61	15669,06	sezon 2	tak	tak	tak	2920,94	1,39	0,22
3	24	10127	422	454	15	1408	1,14	13172,67	sezon 2	tak	tak	tak	1787,33	1,48	0,18
4	30	13750	516	581	15	1887	1,45	16232,06	sezon 2	tak	tak	tak	9537,95	1,87	0,69
5	30	9470	350	427	15	1697	1,31	14113,56	sezon 2	tak	tak	tak	3686,44	1,88	0,39
6	28	10527	376	442	15	1556	1,64	16041,24	sezon 2	tak	tak	tak	4178,76	1,92	0,40
7	17	10700	684	784	15	1588	1,22	20263,87	sezon 2	tak	tak	tak	-7413,87	1,20	-0,69
8	31	16362	528	602	15	1765	1,36	19235,94	sezon 2	tak	tak	tak	4514,06	1,45	0,28
9	17	10700	684	784	15	1349	1,23	20120,94	sezon 2	tak	tak	tak	-9020,94	1,04	-0,84
10	21	7824	447	484	13	743	0,99	9727,43	sezon 2	tak	tak	tak	722,57	1,34	0,09
11	21	9632	459	491	14	1114	1,11	11636,10	sezon 2	tak	tak	tak	1593,90	1,37	0,17
12	23	10560	515	635	14	1190	1,04	11430,36	sezon 2	tak	tak	tak	3319,64	1,40	0,31
13	22	6990	350	430	14	1162	1,01	9105,71	sezon 2	tak	tak	tak	3794,29	1,85	0,54
14	21	8297	395	459	14	1263	1,05	10928,23	sezon 2	tak	tak	tak	1541,78	1,50	0,19
15	21	9090	433	497	14	1036	1,04	10968,52	sezon 2	tak	tak	tak	1881,48	1,41	0,21
16	21	8000	419	520	14	771	1,03	9847,90	sezon 2	tak	tak	tak	952,10	1,35	0,12
17	31	11760	380	445	15	1454	1,32	15814,58	sezon 1	tak	tak	tak	4005,42	1,69	0,34
18	26	8600	331	378	15	1561	1,25	11779,84	sezon 1	tak	tak	nie	5200,16	1,97	0,60
19	24	10570	441	538	15	1621	1,32	14242,71	sezon 1	tak	tak	tak	757,29	1,42	0,07
20	25	10980	440	503	15	1686	1,37	13236,73	sezon 1	tak	tak	nie	5363,27	1,69	0,49
21	31	13230	471	552	15	1802	1,39	15843,46	sezon 1	tak	tak	tak	8356,54	1,83	0,63
22	30	11520	384	472	15	1620	1,25	16784,69	sezon 1	tak	tak	tak	2965,31	1,71	0,26
23	31	14500	468	493	15	1642	1,26	16733,81	sezon 1	tak	tak	tak	1016,19	1,22	0,07
24	31	14400	465	502	15	1513	1,16	15500,69	sezon 1	tak	tak	tak	8099,31	1,64	0,56
25	22	11290	514	571	15	1563	1,20	12614,45	sezon 1	tak	tak	tak	865,55	1,19	0,08
26	30	14930	498	603	15	1645	1,73	16990,02	sezon 1	tak	tak	tak	4509,98	1,44	0,30
27	31	17810	575	613	15	1558	1,64	19222,32	sezon 1	tak	tak	tak	-672,32	1,04	-0,04
28	30	12860	429	530	15	1645	1,73	17773,67	sezon 1	tak	tak	tak	5766,33	1,83	0,45
29	30	17040	569	674	15	1807	1,39	19713,65	sezon 1	tak	tak	tak	1286,35	1,23	0,08
30	12	6850	571	663	15	1779	1,37	19805,28	sezon 1	tak	tak	tak	-11055,28	1,28	-1,61
31	31	17980	580	618	15	1584	1,22	20049,37	sezon 1	tak	tak	tak	-2049,37	1,00	-0,11
32	25	14400	576	680	15	1596	1,45	14674,60	sezon 1	tak	tak	tak	4325,40	1,32	0,30
33	31	10936	423	488	15	1597	1,42	15795,00	sezon 3	tak	tak	tak	5475,83	2,05	0,50
34	30	10304	405	466	15	1588	1,38	14079,57	sezon 3	tak	tak	tak	4563,86	1,77	0,44
35	29	8372	358	413	15	1679	1,41	12553,19	sezon 3	nie	nie	nie	3348,62	1,94	0,40
36	26	11752	503	598	15	1536	1,34	13012,55	sezon 3	tak	tak	tak	2188,21	1,33	0,19
37	26	10224	430	531	15	1430	1,30	13012,55	sezon 3	tak	tak	tak	3282,31	1,53	0,32
38	25	11422	472	548	15	1609	1,39	15104,38	sezon 3	tak	tak	tak	2220,28	1,56	0,19
39	25	8985	389	486	15	1473	1,34	12577,23	sezon 3	tak	tak	tak	1037,22	1,54	0,12
40	24	9015	391	488	15	1476	1,34	12564,77	sezon 3	nie	nie	nie	1176,01	1,54	0,13
41	23	10850	435	500	15	1614	1,39	14725,20	sezon 3	nie	nie	nie	3184,18	1,68	0,29
42	22	8902	441	524	14	1101	1,08	10680,49	sezon 3	nie	nie	nie	1885,16	1,45	0,21
43	21	8784	435	523	14	1112	1,09	10635,87	sezon 3	nie	nie	nie	2458,52	1,46	0,28
44	20	8761	434	523	14	1112	1,09	10628,81	sezon 3	nie	nie	nie	1839,10	1,47	0,21
45	23	11580	475	547	15	1646	1,40	15622,78	sezon 3	nie	nie	nie	2371,98	1,57	0,20
46	21	9126	434	518	14	1158	1,15	11411,48	sezon 3	nie	nie	nie	1562,54	1,46	0,17
47	22	8903	437	530	14	1066	1,10	10670,73	sezon 3	nie	nie	nie	1702,55	1,43	0,19
48	24	9126	409	502	15	1461	1,24	12763,30	sezon 3	tak	tak	tak	801,64	1,47	0,09

Na podstawie globalnej analizy wrażliwości i otrzymanych wyników (tabela 5) można stwierdzić, iż przychód wzrasta wraz ze wzrostem liczby dni pracy pojazdu. Przychód względny i zysk względny został poniżej obliczony na dwa sposoby: uwzględniając wszystkie

warianty zleceń wraz z tymi, które generują stratę oraz uwzględniając tylko warianty przynoszące zysk z realizacji zlecenia.



Rys. 13. Wyniki obliczeń przychodu i zysku względnego w funkcji liczby dni pracy pojazdu

Badania wykazały też, że istotnym czynnikiem wpływającym na przebieg zależności pomiędzy liczbą dni pracy, a przychodem lub zyskiem jest również pora roku. Sezon determinuje tempo wzrostu przychodu względnego i zysku względnego. Tempo to jest wysokie w przedziale 10-22 dni pracy pojazdów i staje się umiarkowane w przedziale 23-25 dni pracy w miesiącu. Natomiast zwiększanie liczby dni pracy powyżej 26-27 nie daje już przyrostu korzyści. Wyniki przedstawione na rys. 13c) potwierdzają, że w celu osiągnięcia zysku względnego pojazdy powinny być użytkowane nie mniej niż 20 dni w miesiącu, natomiast analiza przewidywań bez uwzględniania zleceń powodujących straty (rys. 13d) potwierdza, że jest to powyżej 21 dni pracy dla sezonu 2 oraz 22 dni dla sezonu 1.

6. **Podsumowanie**

Otrzymane i przedstawione w artykule wyniki pozwoliły na następujące stwierdzenia, iż przyjęte miary efektywności ekonomicznej pokazały wpływ liczby dni pracy pojazdów na przychód i zysk z usług transportowych oraz że opracowany model jest przydatny do predykcji miesięcznego przychodu z usług przewozowych.

Wyniki obliczeń dają podstawę do stwierdzenia, że zwiększanie liczby dni pracy pojazdu ma ograniczony wpływ na proces narastania przychodu w firmie. Obserwuje się, że dodatnie wartości przychodu są osiągane przy liczbie dni pracy powyżej 19-20.

Zarówno liczba, jak i rodzaj danych, wykorzystanych w sieci neuronowej pozwoliły na osiągnięcie wysokich wyników analiz, na poziomie 80-90% skuteczności.

Uzyskane wyniki obliczeń pokazały specyficzne cechy wpływu liczby dni pracy na przychód w firmie transportowej. Wzrost liczby dni pracy sprzyja wzrostowi przychodu w sposób ograniczony, a to ograniczenie zależy od pory roku.

Opracowany model sieci neuronowej umożliwia wspomaganie podejmowania decyzji w realizacji procesów transportowych, uwzględniających efektywność ekonomiczną procesu eksploatacji pojazdów samochodowych. Tym samym uzyskane rezultaty pokazały przydatność przyjętych miar efektywności ekonomicznej oraz zbudowanego modelu do predykcji (przewidywania) rezultatów ekonomicznych działalności transportowej firmy.

Literatura

[1] Aleksandrowicz P, Żółtowski B. Vehicle repair costs calculation systems. Polish Association for Knowledge Management. Warszawa: BEL Studio sp. z o.o., 2010.

[2] Andrzejczak K, Młyńczak M, Selech J. Poisson-distributed failures in the predicting of the cost of corrective maintenance. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (4): 602-609, https://doi.org/10.17531/ein.2018.4.11.

[3] Biesok G. Logistyka Usług. Warszawa: CeDeWu, 2013.

[4] Bronk H. Cechy i układ kosztów w transporcie umożliwiające podejmowanie decyzji. Koszty i ceny w transporcie. Pomiar analiza. Szczecin: Zeszyty naukowe 813, 2014; 21 – 38.

[5] Chen X, Xiao L, Zhang X, Xiao W, Li J. An integrated model of production scheduling and maintenance planning under imperfect preventive maintenance. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2015; 17 (1): 70–79.

[6] Chłopek Z, Bebkiewicz K. Model of the structure of motor vehicles for the criterion of the technical level on account of pollutant emission. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (4): 501-507, http://dx.doi.org/10.17531/ein.2017.4.2.

[7] Coupek D, Gulec A, Lechler A,Verl A. Selective rotor assembly using fuzzy logic in the production of electric drives. Conference on Intelligent Computation in Manufacturing Engineering, 2014.

[8] Gohari A, Matori N A, Yusof K W, Toloue I, Sholagberu A T. The effect of fuel price increase on transport cost of container transport vehicles. International Journal of GEOMATE 2018; 15: 174-181.

[9] Himanen V, Nijkamp P, Reggiani A. Neural networks in transport applications. Ashgate 1998. Reissued 2018 by Routledge.

[10] Jóźwiak A. Application of Kohonen's Network in Logistics. Gospodarka Materiałowa i Logistyka 2017; 5: 258-271.

[11] Kijek M, Brzeziński M, Gontarczyk M, Rykała Ł, Zelkowski J. Fuzzy Modeling of Evaluation Logistic Systems. Transport Means 2017; 2: 377-382.

[12] Kleiner F, Friedrich H E. Development of a Transport Application based Cost Modelfor the assessment of future commercial vehicle concepts. Geneva: European Battery, Hybrid and Fuel Cell Electric Vehicle Congress, 2017.

[13] Kowalski M, Magott J, Nowakowski T, Werbińska-Wojciechowska T. Analysis of transportation system with the use of Petri nets Eksploatacja i Niezawodnosc – Maintenance and Reliability 2011; 1 (49): 48–62.

[14] Koźlak A. Ekonomika transportu. Teoria i praktyka. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego, 2010.

[15] Mendyk E. Ekonomika transportu. Poznań: Wyższa Szkoła Logistyki, 2009.

[16] Niewczas A, Rymarz J, Debicka E. Stages of operating vehicles with respect to operational efficiency using city buses as an example. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (1): 21–27, http://dx.doi.org/10.17531/ein.2019.1.3.

[17] Oziemski S. Efektywność eksploatacji maszyn. Podstawy techniczno – ekonomiczne. Radom: Biblioteka problemów eksploatacji, 1999.
[18] Prochowski L. Evaluation of the process of mileage growth during the operation of motor trucks, in several categories of engine cubic capacity. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (3): 359–370, http://dx.doi.org/10.17531/ein.2018.3.3.
[19] Raporty z Internetu Głównego Inspektoratu Transportu Drogowego.
[20] Rudyk T, Szczepański E, Jacyna M. Safety factor in the sustainable fleet management model. Archives of Transport 2019; 49: 103-114.
[21] Świderski A, Borucka A, Jacyna-Gołda I, Szczepański E. Wear of brake system components in various operating conditions of vehicle in the transport company. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (1): 1–9, http://dx.doi.org/10.17531/ein.2019.1.1.
[22] Świderski A, Jóźwiak A, Jachimowski R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 292–299, http://dx.doi.org/10.17531/ein.2018.2.16.
[23] Świderski A. Studies and quality assurance neural modelling of the technical transport means. Archive of Transport. Polish Academy of Sciences Committee of Transport 2009; 21 (3-4).
[24] Teodorovic D, Vukadinovic K. Traffic Control and Transport Planning:: A Fuzzy Sets and Neural Networks Approach. Springer Science+Business Media 2012.
[25] Urbanyi-Popiołek I. Ekonomiczne i organizacyjne aspekty transportu. Bydgoszcz: Wyższa Szkoła Gospodarki, 2013.
[26] Witkowski K, Tanona K. Analiza kosztów transportu drogowego. Logistyka 2013; 5: 411 – 416.
[27] Zał. nr 2. do U. z dnia 20 czerwca 1997 r. Prawo o ruchu drogowym.

Analysis of the impact of the use time of N1 motor vehicles on the economic efficiency of their maintenance

Keywords: *operation of vehicles, motor vehicles of category N1, economic efficiency, neural networks*

Abstract: The efficiency of operation of motor vehicles with a DMC (Permissible Laden Mass) <3.5 tonnes is considered. These are vehicles belonging motor vehicles of category N1, usually referred to as delivery vehicles. The results of observations on the implementation of transport orders in 7 transport companies from the MŚP (Small and Middle-size Companies) sector were used to conduct the effectiveness analysis. The research group covered 24 vehicles that implementation transport orders in the urban zone and in the immediate vicinity of the city. Information was collected on a monthly basis.

During the analysis of economic efficiency the income measures (absolute and relative) were used. The calculations were carried out using the model of the vehicle operation process in the form of a neural network, in which a set of 12 input variables and 3 output variables were taken into account. Using the Statistica 13.3 computer program and defining the group and factors describing the process of implementation of individual transport tasks, the developed neural network model enabled searching for the impact of selected operational factors on the economic efficiency of N1 category cars.

The calculations showed a significant impact of the number of vehicle days in a month, the weight of the load, as well as the time of year. The obtained calculation results showed the specific features of the impact of the number of working days on revenue in a transport company. The increase in the number of working days favors the increase in income in a limited way, and this restriction depends, among others since the time of year.

Prof. dr hab. inż. Ireneusz Pielecha

Instytut Silników Spalinowych i Transportu
Politechnika Poznańska
Piotrowo 3, 60-965 Poznań, Polska
E-mail: Ireneusz.Pielecha@put.poznan.pl

Prof. dr hab. inż. Jacek Pielecha

Instytut Silników Spalinowych i Transportu
Politechnika Poznańska
Piotrowo 3, 60-965 Poznań, Polska
E-mail: Jacek.Pielecha@put.poznan.pl

Symulacyjna analiza energochłonności pojazdów elektrycznych w testach badawczych

Słowa kluczowe: pojazdy samochodowe, napęd elektryczny, akumulatory wysokonapięciowe, przepływ energii, testy jezdne

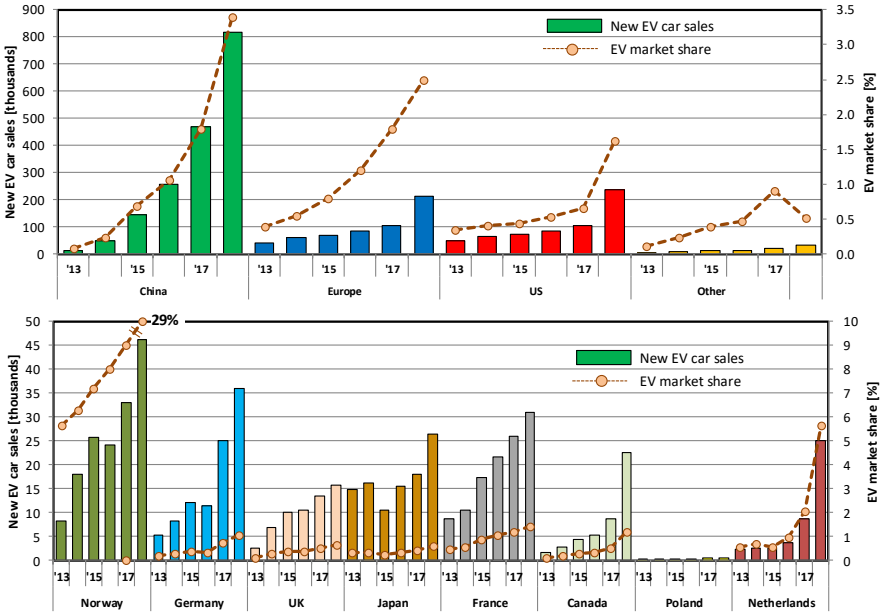
Streszczenie: Ocena przepływu energii przez układy pojazdów elektrycznych umożliwia oszacowanie ich energochłonności. W artykule przedstawiono analizy dotyczące zużycia energii pojazdów elektrycznych w wybranych testach jezdnych (NEDC, WLTC oraz w rzeczywistych warunkach ruchu – test RDC) w odniesieniu do zróżnicowanej masy pojazdów. Analizie poddano również wykorzystanie silników elektrycznych, przedstawiając mapy ich pracy, wielkości przepływu energii w akumulatorach oraz stopień zmiany ich naładowania. Badania i analizy symulacyjne wykonano z wykorzystaniem oprogramowania AVL Cruise. Stwierdzono, że mimo podobnych wartości energochłonności pojazdów w testach badawczych NEDC oraz RDC, to występują znaczące różnice przepływu energii w układach akumulacji pojazdów. Zmiany stopnia naładowania akumulatora odniesione do 100 km testu są zbliżone w testach WLTC oraz RDC (różnica 6%); dla testu NEDC różnica ta wynosi maksymalnie 25% (w odniesieniu do poprzednich testów). Energochłonność pojazdów elektrycznych jest silnie zależne od testu badawczego; wartości uzyskane w testach kształtują się na poziomie 10,1–13,5 kWh/100 km (test NEDC); 13–15 kWh/100 km (test WLTC) oraz 12,5–16,2 kWh/100 km w teście RDC. Wartości energochłonności w testach NEDC oraz WLTC są odpowiednio mniejsze o około 20% i 10% w odniesieniu do testu RDC. Zwiększenie masy pojazdu zwiększa zużycie energii (zwiększenie o 100 kg masy pojazdu zwiększa zużycie energii o 0,34 kWh/100 km).

1. Wprowadzenie

Konieczność ograniczania zużycia paliwa przez pojazdy samochodowe wyposażone w silniki spalinowe (oraz napędy hybrydowe) skutkuje zwiększeniem produkcji pojazdów z elektrycznymi układami napędowymi, co prowadzi do zwiększania ich udziału w ogólnej liczbie pojazdów. Zaostrzenie standardów dotyczących ograniczenia zużycia paliwa z floty pojazdów w 2030 r., powinno przyczynić się do znaczącego rozwoju elektromobilności [3] przy jednoczesnym ograniczeniu emisji gazów cieplarnianych.

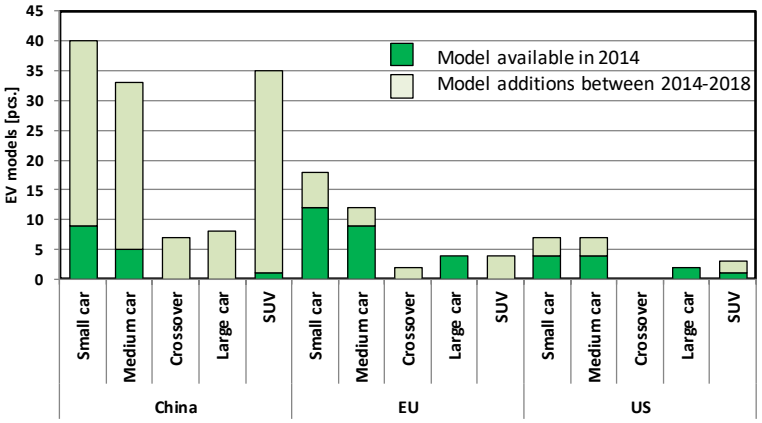
Sprzedaż pojazdów elektrycznych (BEV – battery electric vehicle) na światowych rynkach ulega zwiększeniu, jednak udział tych pojazdów w ogólnym rynku samochodów nadal nie jest zbyt duży. Średni europejski udział pojazdów EV (electric vehicle) w rynku wynosi około 2,5%, mimo sprzedaży tych pojazdów w 2018 r. na poziomie 200 tys. szt.

(rys. 1). Najbardziej dynamicznym rynkiem jest rynek chiński, na którym notuje się ponad 800 000 nowych rejestracji w 2018 r. Mimo tego udział pojazdów BEV wynosi tam około 3,5%. W Europie przoduje Norwegia z 45 tysiącami nowych rejestracji w 2018 r. (udział w rynku pojazdów BEV wynosi obecnie 29%). W Polsce udział pojazdów BEV określany jest na poziomie 0,4% [10]; na koniec lipca 2019 r. zarejestrowanych było 4009 pojazdów elektrycznych [14].



Rys. 1. Światowa sprzedaż pojazdów elektrycznych i ich udział w rynkach wybranych państw i regionów świata [5]

W ostatnich kilku latach obserwuje się duży przyrost modeli pojazdów elektrycznych, w zróżnicowanych segmentach (rys. 2). Największy przyrost modeli dotyczy rynku chińskiego – we wszystkich segmentach. Dostępnych jest tam około 120 modeli samochodów EV. Na rynku europejskim około 20 modeli – i dwukrotnie mniej na rynku amerykańskim [5]. Z ogólnej sprzedaży pojazdów w 2018 r. na rynku chińskim –90% to pojazdy typu small car. Na rynku europejskim obserwuje się rozwój pojazdów elektrycznych typu crossover oraz SUV.



Rys. 2. Modele pojazdów elektrycznych z podziałem na kategorie dostępne na wybranych rynkach świata [5]

Na rynku polskim jest oficjalnie dostępnych 20 modeli pojazdów typu BEV (wliczając w to samochody dostawcze) [14]. Ich zasięg jest dość zróżnicowany i wynosi od 100 km do 540 km (w teście NEDC – New European Driving Cycle), co zapewniają akumulatory o pojemności, odpowiednio, od 6,1 kWh do 90 kWh. Zestawienie nie uwzględnia pojazdów Tesli, gdyż ich oficjalna sprzedaż (on-line) rozpoczęła się dopiero pod koniec sierpnia 2019 r.

Z danych przedstawionych przez Transport & Environment [10] wynika, że udział pojazdów elektrycznych będzie wynosił około 8% całej sprzedaży w 2025 r. oraz około 17% w 2030 r.

Analizę energochłonności napędów konwencjonalnych prowadzi się w oparciu o wartości zużycia paliwa lub emisji dwutlenku węgla. Jednakże analizy emisyjne napędów spalinowych pojazdów samochodowych wskazują na duże rozbieżności (o wartości 30–40%) w zużyciu paliwa i emisji CO_2 w badaniach certyfikacyjnych i w warunkach rzeczywistych [4].

Na wartość emisji dwutlenku węgla mają również wpływ warunki pracy pojazdów oraz wyposażenie ich w systemy ograniczania zużycia paliwa na postoju. Wpływ różnych tras badawczych oraz aktywacji systemu start-stop były przedmiotem badań prowadzonych podczas testów w rzeczywistych warunkach ruchu dla samochodów osobowych [8]. Stwierdzono w nich, że niepowtarzalność warunków pracy pojazdów na tej samej trasie badawczej mogą powodować różnicę w wartości emisji dwutlenku węgla około 26%. Dodatkowo stosowanie systemu start stop umożliwia zmniejszenie tej emisji o kolejne 11–15%.

Wpływ na emisję dwutlenku węgla w testach drogowych oceniano również przez pryzmat warunków dynamicznych w testach drogowych [6]. Testy takie, określane jako RDC (Real Driving Conditions), są wymagane od 2018 roku jako część procesu homologacyjnego pojazdów. Autorzy stwierdzili, że istnieje silna zależność między emisją drogową dwutlenku węgla a warunkami dynamicznymi jazdy. Wykazano, że dwukrotne zwiększenie wartości względnego przyspieszenia dodatniego skutkuje 3-krotnym zwiększeniem emisji drogowej dwutlenku węgla.

Według Pavlovic et al. [11] emisja CO_2 w teście WLTC jest o około 10% większa od analogicznej w teście NEDC (dla pojazdów z silnikami ZI o masie około 1500 kg), natomiast zużycie energii jest większe o około 40% (dla pojazdów wyposażonych w silniki ZI).

Badania symulacyjne zużycia paliwa i emisji CO_2 z wykorzystaniem AVL Cruise prowadzone przez Tsokolis et al. [16] wskazują na różnice wartości uzyskiwanych z testów NEDC oraz WLTC. Różnice te dotyczą ponad 63% badanych przypadków pojazdów z silniami ZI oraz 81% z silnikami ZS. Emisja CO_2 była większa średnio dla wszystkich pojazdów o 11% w teście WLTC w odniesieniu do testu NEDC. Średnie sprawności uzyskano większe w teście WLTC (niż w NEDC) i wynosiły one odpowiednio: sprawności silnika – 31% (do 25%) oraz sprawności pojazdu – 26% (do 21%).

Analiza napędu pojazdu elektrycznego jest obecnie rozpatrywana w dwóch formach. Pierwsza dotyczy badań pojazdów, druga – prowadzenia badań symulacyjnych. Obydwa warianty badań mogą być łączone w celu określenia energochłonności pojazdu elektrycznego.

Analizy prowadzone przez Wu et al. [18] wskazują na zwiększoną sprawność napędu elektrycznego podczas ruchu miejskiego niż w warunkach ruchu autostradowego. Wynika to z większych możliwości odzyskiwania energii w warunkach jazdy miejskiej.

Analizę przepływu energii w pojazdach hybrydowych w rzeczywistych warunkach ruchu drogowego prowadzi się od kilku lat [12, 13, 17]. Analiza dotycząca takich układów napędowych opiera się przede wszystkim na wykorzystaniu elektrycznej analizy energetycznej z pominięciem zużycia paliwa przez pojazdy. Wynika to z faktu pracy silnika spalinowego częściowo jako generatora prądu umożliwiającego zwiększenie pojemności energetycznej akumulatora pojazdu.

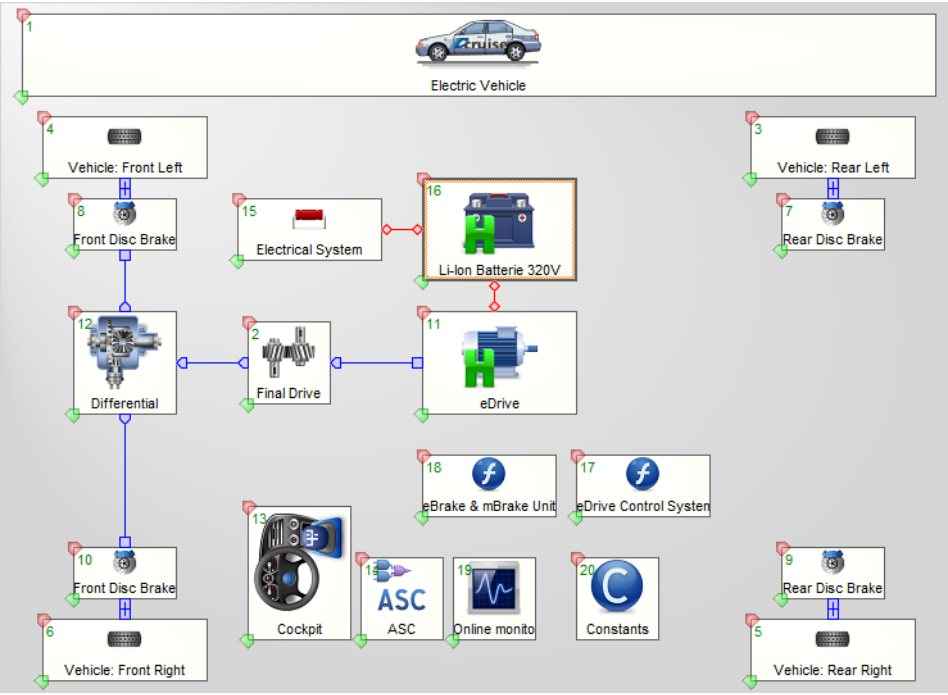
Obecne prace dotyczące redukcji zużycia energii w pojazdach elektrycznych dotyczą analiz związanych z optymalizacją momentu obrotowego silników elektrycznych [19], ograniczaniem strat pracy silników elektrycznych [15], profilem prędkości pojazdów [1], problematyką wyznaczania trasy w aspekcie stacji ładowania [7, 20], a także optymalizacji infrastruktury ich ładowania [2, 9].

Głównym celem badań jest określenie różnic w jednostkowym zużyciu energii elektrycznej przez pojazdy samochodowe w testach badawczych. Konieczna jest zatem odpowiedź na pytanie: w jakim stopniu różne testy wpływają na szacowanie jednostkowego zużycia energii i czy możliwe jest stosowanie zamienne takich testów. Dodatkowym czynnikiem uwzględnianym w badaniach będzie zmienna masa pojazdu.

2. Metodyka badań

Badania prowadzono z wykorzystaniem oprogramowania symulacyjnego AVL Cruise. Jest to system pozwalający na symulację układu napędowego (BEV, HEV lub konwencjonalnego), którego szczególne cechy to: (1) podział układu napędowego na elementy funkcjonalne o predefiniowanych charakterystykach poszczególnych komponentów; (2) niezależność struktury modelu oraz algorytmu rozwiązania; (3) generowanie równań na podstawie pełnej definicji system oraz (4) integralność wielowątkowych modeli danych.

Modelowano pojazd elektryczny o różnej masie własnej, którego model strukturalny przedstawiono na rys. 3. Charakterystykę techniczną pojazdu przedstawiono w tabeli 1.

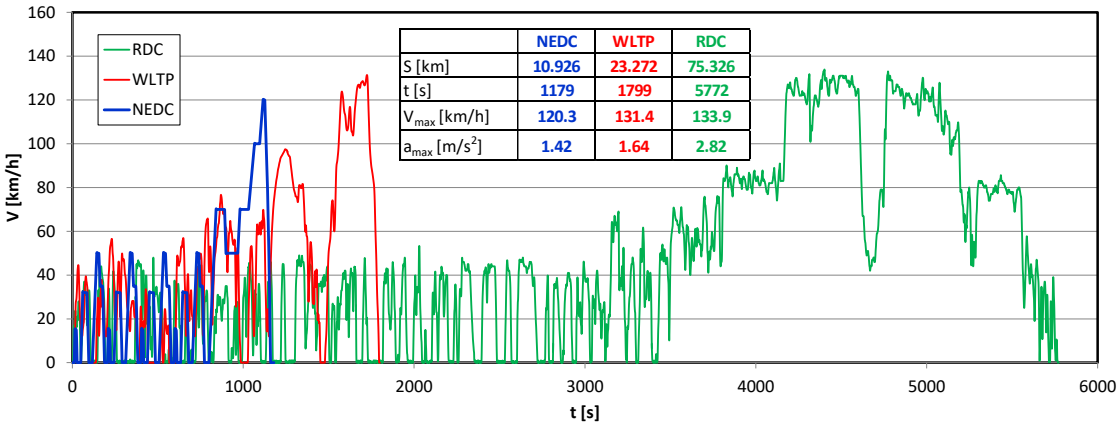


Rys. 3. Schemat układu napędu elektrycznego pojazdu (AVL Cruise)

Tabela 1. Dane techniczne symulowanego pojazdu i napędu

Pojazd		
Masa własna pojazdu	1000; 1500; 2000	kg
Rozstaw osi	2467	mm
Układ elektryczny		
Akumulator	Li-Ion, 25 kWh	
Napięcie nominalne	360	V
Pojemność celi × liczba celi	36 × 2	Ah
Silnik elektryczny		
Typ	asynchroniczny	
Moment obrotowy	240 @ 0–3000	Nm@rpm
Przekładnia		
Przełożenie	6.058	

Badania energochłonności prowadzono w odniesieniu do zróżnicowanych testów jezdnych (NEDC, WLTC i RDC) oraz pojazdu o różnych masach własnych (1000; 1500 oraz 2000 kg). Charakterystykę testów badawczych przedstawiono na rys. 4. W badaniach uwzględniono dotychczasowy test jezdny NEDC, współczesny test badawczy WLTC oraz test drogowy RDC. Cechuje je zróżnicowany profil jazdy, różne długości tras przy zbliżonych wartościach prędkości maksymalnych. Ze względu na różne warunki przyspieszania pojazdu spodziewane są różne wartości całkowitego zużycia energii.



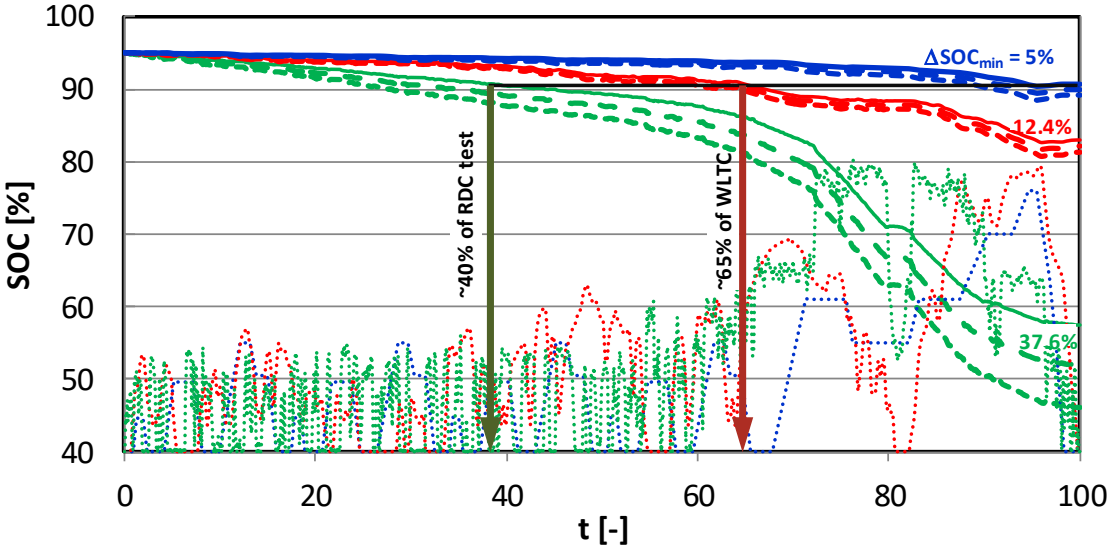
Rys. 4. Testy badawcze wykorzystane w badaniach oraz ich charakterystyki

3. Charakterystyka napędu elektrycznego pojazdu

Badania napędu elektrycznego prowadzono przy stałej wartości początkowej stopnia naładowania akumulatora (SOC = 95%). Ze względu na możliwość doładowania jedynie podczas hamowania regeneracyjnego nie dokonano zmian tej wartości. Stwierdzono, że rodzaj testu badawczego znacząco wpływa na końcową wartość naładowania akumulatora (rys. 5). Najkrótszy test (NEDC) powoduje kilkuprocentową zmianę SOC akumulatora. Zwiększenie długości i intensywności testu skutkuje mniejszymi wartościami końcowymi SOC. Uwzględniając długość testu badawczego określono końcową wartość SOC przypadającą na 100 km testu. W teście NEDC wartość końcowa ΔSOC/100 km (przy uwzględnieniu 95% początkowego naładowania akumulatora) wynosi 53%, w teście WLTC – 42%, a w teście RDC – 45%. Oznacza to, że najbardziej agresywnym teście jezdnym RDC wartości końcowe SOC nie są najmniejsze.

Maksymalne różnice ΔSOC między testami wynoszą 11% (NEDC i WLTC) i są one niewielkie przy porównaniu ΔSOC w testach WLTC oraz RDC – wynoszą 3%.

Dodatkowo zwiększenie masy pojazdu powoduje 5% zmianę SOC (test NEDC), 12-procentową zmianę podczas testu WLTC. Podczas testu RDC masa pojazdu znacząco wpływa na zmianę końcowych wartości SOC. Zwiększenie masy o 100 kg zwiększa wyładowanie akumulatora średnio o około 4%. Jest to wartość bardzo istotna, gdyż przy minimalnej masie pojazdu (1000 kg) zmiana wartości SOC w tym teście wynosi aż 95 – 57 = 38%. Biorąc pod uwagę, że energia całkowita akumulatora wynosi 25 kWh oraz jego możliwości pracy w przedziale rzeczywistego rozładowania wynoszące 20–80%, to zmiana SOC wynosi już 63% możliwej do wykorzystania energii akumulatora. Oznacza to, że 1/3 zmian SOC odpowiada za 2/3 skutecznej energii akumulatora Li-Ion.



Rys. 5. Analiza zmian stopnia naładowania akumulatorów dla różnych mas pojazdu (1000 kg – linia ciągła, 1500 kg – linia przerywana o małej częstości, 2000 kg – linia o dużej częstości) i różnych testów badawczych

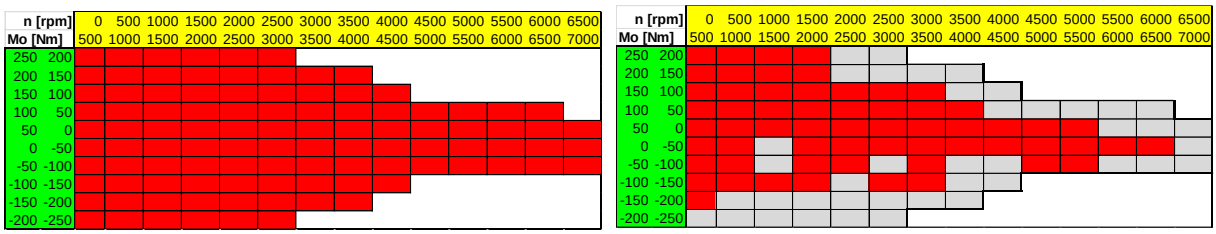
Dodatkowo stwierdzono, że maksymalna zmiana SOC w teście NEDC występuje po 40% długości testu RDC oraz po 65% czasu trwania testu WLTC. Oznacza to duży wpływ jakości testu na końcową wartość SOC akumulatora.

Do oceny stopnia wykorzystania pola pracy silników elektrycznych wykorzystano stałą matrycę we współrzędnych Mo = f(n) przy następujących założeniach (uproszczony schemat przedstawiono na rys. 6a):

- n = 250 obr/min – co skutkuje 30 przedziałami (w zakresie 0–7000 obr/min),
- Mo = 20 Nm – co skutkuje 25 przedziałami (w zakresie –250–250 Nm).

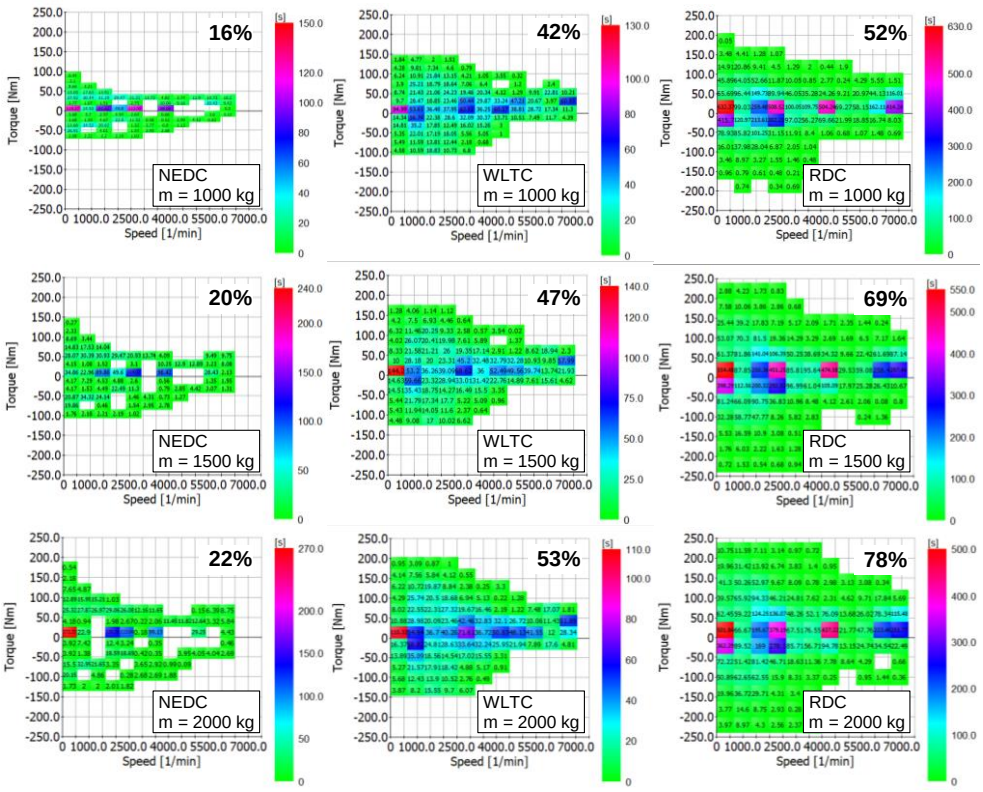
Przykładową mapę wykorzystania silnika elektrycznego w teście RDC dla masy pojazdu m = 2000 kg przedstawiono na rys. 6b. Widoczne jest duże zagęszczenie maksymalnych wartości momentu obrotowego przy małych prędkościach obrotowych, oraz niewykorzystane obszary o dużej sprawności silnika. Hamowanie regeneracyjne pozwala na wykorzystanie również dużego pola pracy (praca dwukwadrantowa silnika elektrycznego), przy czym wartości maksymalne są ograniczone szczególnie w środkowym zakresie prędkości obrotowej tego silnika.

a) b)



Rys. 6. Przykładowa metodyka wyznaczania obszarów użytecznych pola pracy silnika elektrycznego: a) pełne pole pracy, b) rzeczywiste pole pracy silnika elektrycznego w teście RDC

Dla tak wykonanych map ustalono stopień wykorzystania pełnej charakterystyki pracy silnika elektrycznego. Wartości wykorzystania pola pracy zawarto na rysunku 7. W każdym jednostkowym polu zawarto informację o czasie wykorzystania danego obszaru o wartościach ΔMo i Δn. Jednocześnie wartości te są zgodne z legendą przedstawiona przy każdej charakterystyce. Natomiast pojedyncze wartości procentowe wskazują na wielkość wykorzystania maksymalnego (całkowitego) dostępnego pola pracy silnika elektrycznego. Charakterystyczne jest zwiększenie wykorzystania pola pracy silnika elektrycznego przy zmianie testu jezdnego. Dla testu NEDC wykorzystanie pola pracy wynosi maksymalnie 22%, natomiast dla testu RDC – 78% (przy masie m = 2000 kg). Zwiększenie momentu obrotowego silnika elektrycznego jest najistotniejsze przy zwiększaniu masy pojazdu w teście symulującym rzeczywiste warunki jazdy (test RDC).



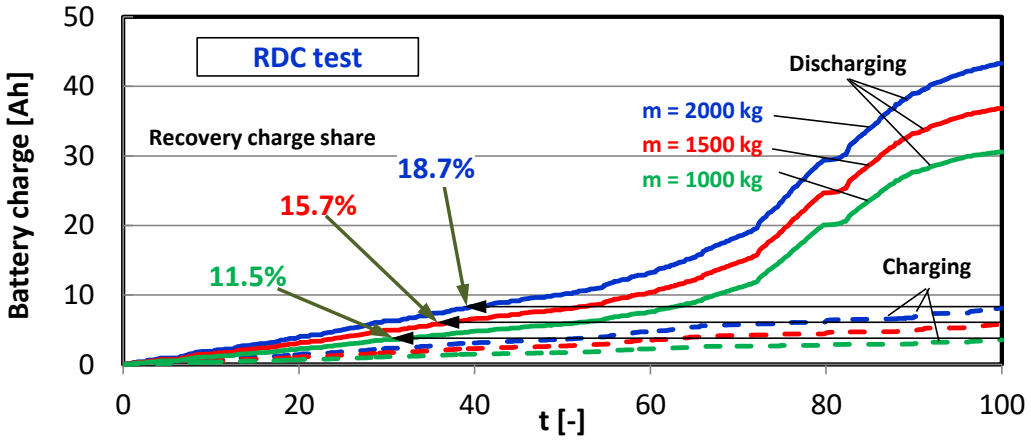
Rys. 7. Charakterystyki pracy silnika elektrycznego w różnych testach badawczych i przy różnych masach pojazdu

4. Analiza przepływu energii przez akumulator

Akumulator wysokonapięciowy o napięciu nominalnym 360 V zasymulowano jako układ składający się z dwóch rzędów cel o pojemności elektrycznej 36 Ah. Taki układ powoduje, że całkowita wartość zgromadzonej energii wynosi około 25 kWh.

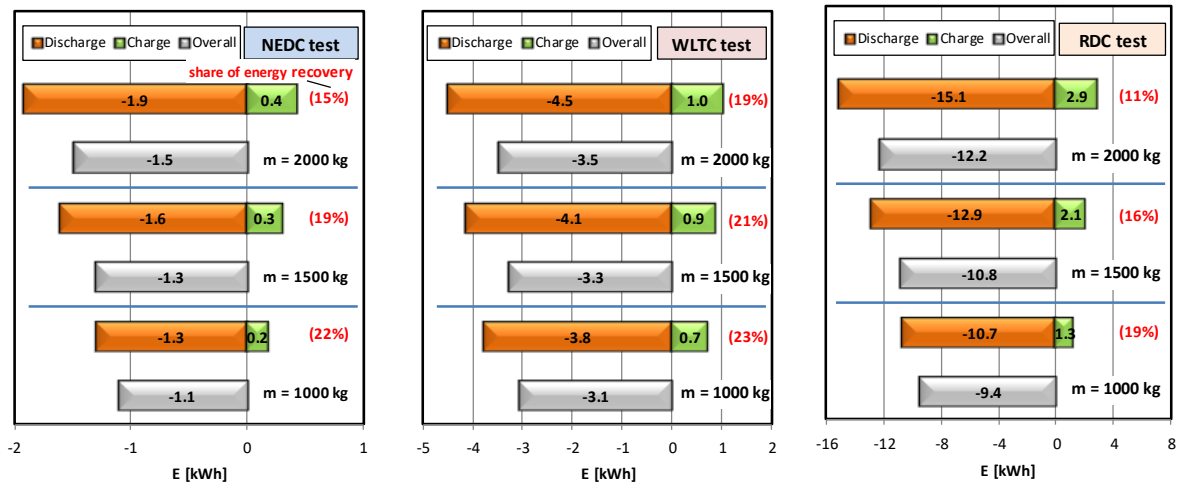
Symulacji przepływu energii dokonano dla wszystkich tras przejazdu, natomiast przykładowy przebieg ładowania i rozładowania akumulatora zawarto na rys. 8. Ze względu na duże wartości zmian SOC (rys. 5) wyniki przedstawiono dla testu RDC. Z przedstawionych danych wynika, że możliwy jest odzysk jedynie kilkunastu procent ładunku akumulatora do pokonania trasy przejazdu.

Mimo, że test RDC generuje największe zmiany SOC, to nie przyczynia się do odzyskiwania dużych względnych wartości ładunku. W odniesieniu do testów NEDC oraz WLTC średnie wartości udziału ładowania są o około 3–4 punkty procentowe większe dla każdej masy pojazdu.



Rys. 8. Zmiany ładunku akumulatora podczas testu RDC przy różnych masach pojazdu oraz charakterystyki jego ładowania i rozładowania

Szczegółowe dane dotyczące przepływu energii przedstawiono na rys. 9. Określono tam warunki ładowania i rozładowania akumulatorów oraz całkowitą zmianę ich energii. Z zamieszczonych danych wynika, że zwiększenie masy pojazdu, mimo zwiększenia odzyskiwanej energii, wpływa na ograniczenie jej udziału w całkowitym bilansie. Taki wynik wskazuje na konieczność optymalizacji masy pojazdu w kontekście przepływu energii w pojeździe elektrycznym. Duże przepływy energii podczas testu RDC wskazują również na konieczność zapewnienia dużych pojemności energetycznych akumulatorów pojazdów BEV oraz dużych początkowych wartości SOC. Wartości całkowitej zmiany energii są największe podczas testu RDC i są 3–4-krotnie większe od zmian energii w teście WLTC. Różnice między zmianami energii w testach NEDC oraz RDC są około 8-krotne (na korzyść testu NEDC). Zwiększenie masy pojazdów (o 1/3) wpływa w około 12% na ogólne zwiększenie zużycia energii, niezależnie od testu. Najmniejszy wpływ masy na zużycie energii uzyskano w teście NEDC (około 6%), największy w teście RDC – 17%.



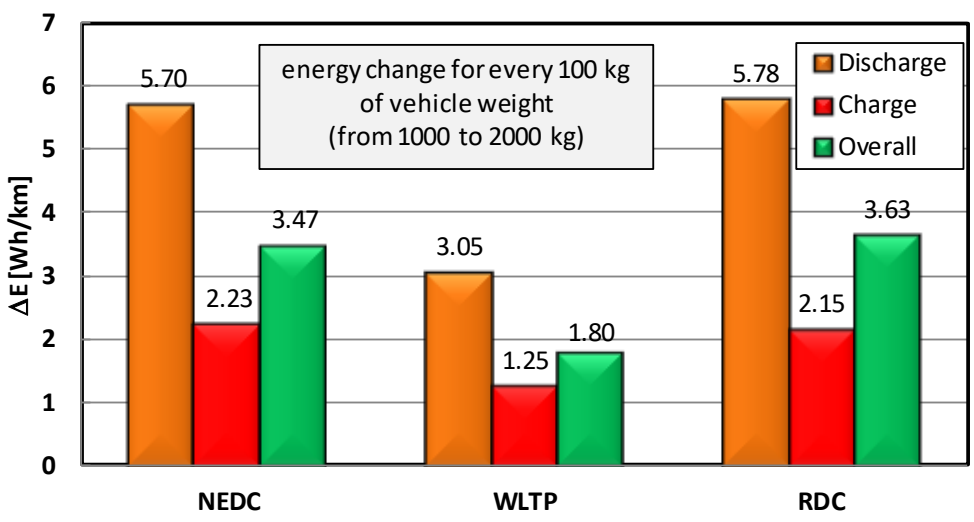
Rys. 9. Warunki pracy akumulatorów i przepływu energii podczas testów jezdnych

Ocenę zapotrzebowania energetycznego w postaci energii odniesionej do 100 kg masy pojazdu oraz do długości testu jezdnego określono równaniem:

$$\Delta E = \frac{E_{i,m_2} - E_{i,m_1}}{n \cdot S_{test,j}} \tag{1}$$

gdzie: E – wartość energii akumulatora [Wh]; i – fazy jazdy uwzględniające rozładowanie, ładowanie oraz całkowity przepływ energii; j – realizowane testy jezdne: NEDC, WLTC, RDC, S – dystans testu jezdnego [km]; m_2 = 2000 kg, m_1 = 1000 kg; n = 10 – oznacza odniesienie energii do 100 kg masy pojazdu.

Wyniki tych obliczeń przedstawiono na rys. 10. Z analizy wynika, że podobne nakłady energetyczne na 100 kg masy pojazdu oraz na jednostkową długość testu występują dla testów NEDC oraz RDC. Mimo, że test NEDC w odniesieniu do WLTC jest mniej dynamiczny, to jednostkowe nakłady energetyczne zgodnie z równaniem (2) są znacznie większe (o około 50%). Należy zauważyć, że nakłady takie są zbliżone w testach NEDC oraz RDC; podobieństwa te występują zarówno podczas rozładowania oraz ładowania akumulatorów.



Rys. 10. Analiza zmian przepływu energii przez akumulator przy uwzględnieniu zmiany masy pojazdu o wartości Δm = 100 kg

5. Analiza jednostkowego zużycia energii pojazdu

Zużycie energii pojazdu określono w postaci energii jednostkowej dla każdego z testów jezdnych oraz przy uwzględnieniu masy pojazdu w zakresie od 1000 kg do 2000 kg. Jednostkowe zużycie energii (energochłonność) [kWh/100 km] określono jako wartość zużycia energii odniesione do długości testu o wartości 100 km:

$$E_j = \frac{\int_{t=0}^{t} E dt}{S} \cdot 100 \tag{2}$$

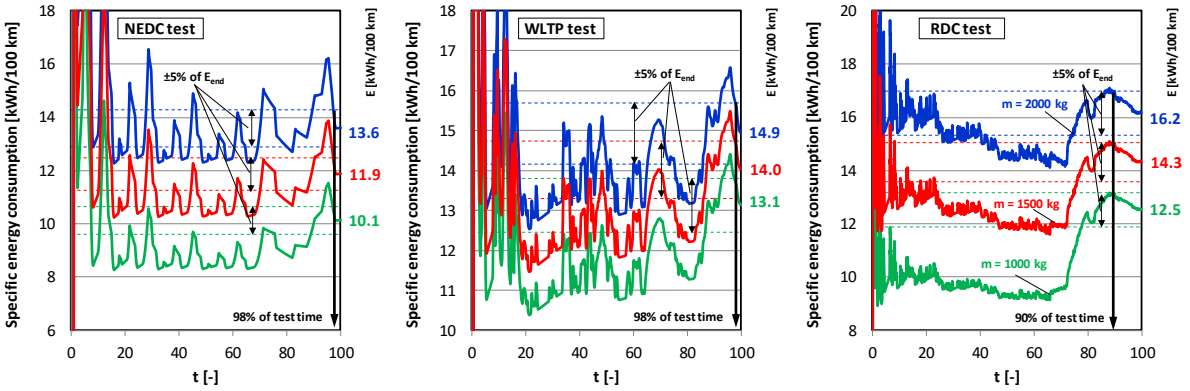
gdzie: t_0 – oznacza początek testu, t_{max} – bieżący czas testu, S – długość testu a E – bieżącą wartość zużytej energii w teście.

Energię tę określono na dwa sposoby: w całym teście (E_{end}) oraz jako funkcję bieżącego czasu trwania testu – E_j (rys. 11). Początkowa faza każdego z testów generuje duże wartości E_j (ze względu na krótki czas trwania testu), natomiast w dalszej części – następuje stabilizacja tej wartości. Od dynamiki testu badawczego zależy, czy końcowa wartość energochłonności w teście jest stabilizowana w dłuższym czy krótszym okresie czasu. Specyfika testów jezdnych powoduje, że nawet podczas statycznego testu NEDC wartość końcowa energochłonności jest uzyskiwana po 98% czasu jego trwania. Wartość tę ustalono jako przedział 5-procentowych zmian wartości ustalonej od wartości końcowej:

$$t_{\%} = \pm 5\% \text{ of } E_{end} \tag{3}$$

gdzie: $t_{\%}$ – względny czas stabilizacji (w przedziale ±5% ustalonej końcowej wartości energii całkowitej – E_{end}).

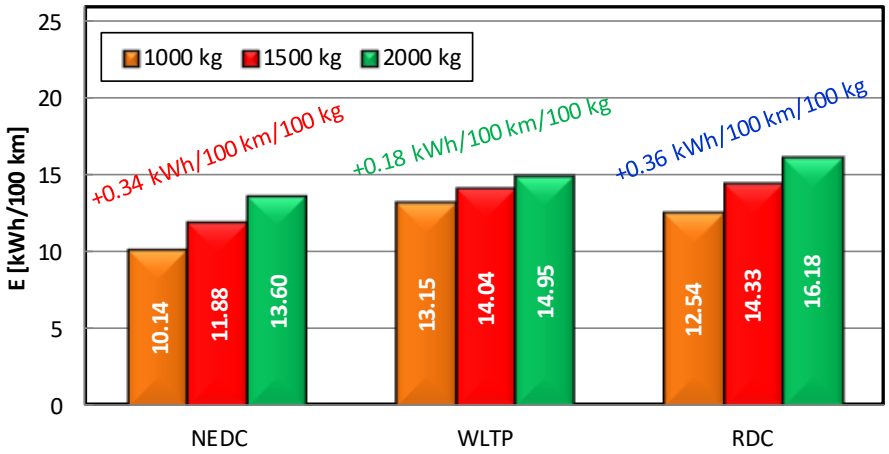
Podczas analiz bardziej dynamicznych testów WLTC oraz RDC, wartości $t_{\%}$ stanowiły odpowiednio: 98% oraz 90% czasu trwania testu (rys. 11a). Oznacza to, że dla najbardziej dynamicznego testu RDC czas ustalenia końcowej wartości całkowitego zużycia energii jest najkrótszy (rys. 11c). Należy zauważyć, że modyfikacja faz testu (jazda miejska, pozamiejska, autostradowa) może przyczynić się do wcześniejszego ustalenia wartości końcowej (przy założeniu niezmienności kolejności faz jazdy).



Rys. 11. Warunki określania końcowej energochłonności napędu podczas testów badawczych

Najkrótszy czas stabilizacji energochłonności uzyskiwany jest podczas testu RDC. W tym teście (najbardziej dynamicznym) wartość końcowa zużycia energii jest również największa. Najmniejsze wartości zużycia tej energii występują w teście NEDC. Należy stwierdzić proporcjonalny wpływ masy pojazdu na wartości końcowe energochłonności każdego z testów.

Wyniki analiz całkowitego zużycia energii przedstawiono na rys. 12. Największe wartości tego zużycia uzyskano podczas testu RDC. Są one większe o około 20% od wartości energochłonności w teście NEDC. Zaznaczono tam również jednostkowe przyrosty tej energochłonności odniesione do zwiększenia masy pojazdu o 100 kg. Stwierdzono, że przyrosty zużycia energii są jednakowe dla testów NEDC oraz RDC i wynoszą one około 0,34 kWh/100 km na każde dodatkowe 100 kg masy pojazdu. Odnotowano dwukrotnie mniejsze przyrosty zużycia energii uzależnione od masy pojazdy podczas testu WLTC.



Rys. 12. Wartości energochłonności pojazdu w testach badawczych przy uwzględnieniu zmiany masy pojazdu

Takie szacowanie wartości przyrostu energochłonności pozwala na określenie tego zużycia bez konieczności dokonywania symulacji lub rzeczywistych badań pojazdu przy zmienionej jego masie własnej.

6. Wnioski

Na podstawie badań symulacyjnych (wykorzystując oprogramowanie AVL Cruise) określono warunki pracy silników elektrycznych pojazdu EV oraz energochłonność pojazdu o zróżnicowanych masach własnych w różnych testach badawczych. Na ich podstawie sformułowano następujące wnioski szczegółowe:

1. W odniesieniu do poziomu naładowania akumulatorów:
   - Dynamika testu RDC powoduje największe zmiany stopnia naładowania akumulatora (ΔSOC). Są one 3-krotnie większe niż w odniesieniu do testu WLTC oraz 7-krotnie większe w odniesieniu do testu NEDC.
   - Uwzględnienie długości testu powoduje, że stosunek zmian ΔSOC na 100 km testu odpowiednio RDC:WLTC:NEDC jest następujący: 1,06:1:1,25. Oznacza to zmianę ΔSOC/100 km w wartościach procentowych odpowiednio: 53%:42%:45%. Najmniejsze różnice odnotowano w teście WLTC oraz RDC – tylko 6%.
2. W odniesieniu do przepływu energii przez akumulator:
   - Wartości całkowitej zmiany energii ładowania i rozładowania akumulatora są największe podczas testu RDC; są one 3–4-krotnie większe od zmian energii w teście WLTC. Różnice między zmianami energii w testach NEDC oraz RDC są około 8-krotne (na korzyść testu NEDC).

- Zwiększenie masy pojazdów (o 1/3) wpływa w około 12% na ogólne zwiększenie zużycia energii (energochłonność), niezależnie od testu. Najmniejszy wpływ masy na całkowity przepływ energii przez akumulator uzyskano w teście NEDC (około 6%), największy w teście RDC – 17%.

3. W odniesieniu do zużycia energii (energochłonności) w teście:
   - Największe wartości zużycia energii uzyskano podczas testu RDC. Są one większe o około 20% od wartości zużycia energii w teście NEDC. Podczas testu WLTC wartości energochłonności są większe od 30 do 10% od energochłonności w teście NEDC (odpowiednio do zwiększania masy pojazdu).
   - Zwiększenie energochłonności w testach NEDC oraz RDC jest jednakowe i wynosi około 0,34 kWh/100 km na każde dodatkowe 100 kg masy pojazdu. Odnotowano dwukrotnie mniejsze przyrosty energochłonności uzależnione od masy pojazdy podczas testu WLTC.

Uzyskane wyniki badań i analiz zużycia energii pojazdów elektrycznych wskazują na konieczność prowadzenia dalszych prac symulacyjnych oraz badań w rzeczywistych warunkach ruchu. Zużycie energii w takich pojazdach jest silnie zależne od testu badawczego; wartości uzyskane w testach badawczych kształtują się na poziomie 10,1–13,5 kWh/100 km (test NEDC); 13–15 kWh/100 km (test WLTC) oraz 12,5–16,2 kWh/100 km w teście RDC. Różnice w ustalonych testach badawczych (NEDC oraz WLTC) wynoszą do 25% (większe w teście WLTC dla pojazdów o mniejszej masie własnej). Badania prowadzone w warunkach rzeczywistych wykazują podobne wartości zużycia energii (pojazdy lekkie – około 1000 kg) oraz zwiększenie tego zużycia o kolejne 10% w odniesieniu do pojazdów o masie około 2000 kg.

Podziękowania
Badania były prowadzone z wykorzystaniem oprogramowania AVL Cruise dzięki programowi AVL University Partnership Program.

Bibliografia

1. Basso R, Kulcsár B, Egardt B, Lindroth P, Sanchez-Diaz I. Energy consumption estimation integrated into the Electric Vehicle Routing Problem. Transportation Research Part D: Transport and Environment 2019; 69: 141–167, https://doi.org/10.1016/j.trd.2019.01.006.

2. Davidov S, Pantoš M. Planning of electric vehicle infrastructure based on charging reliability and quality of service. Energy 2017; 118: 1156–1167, https://doi.org/10.1016/j.energy.2016.10.142.

3. European Commission. Proposal for a regulation of the European Parliament and of the Council setting emission performance standards for new passenger cars and for new light-commercial vehicles as part of the Union's integrated approach to reduce CO2 emissions from LDVs. Brussels, 8.11.2017, SWD(2017) 650 final. ec.europa.eu/clima/sites/clima/files/transport/vehicles/docs/swd_2017_650_p1_en.pdf.

4. Fontaras G, Zacharof N-G, Ciuffo B. Fuel consumption and CO_2 emissions from passenger cars in Europe – Laboratory versus real-world emissions. Progress in Energy and Combustion Science 2017; 60: 97–131, https://doi.org/10.1016/j.pecs.2016.12.004.

5. IEA, Global EV Outlook 2019. IEA, Paris, www.iea.org/publications/reports/globalevoutlook2019.

6. Kurtyka K, Pielecha J. The evaluation of exhaust emission in RDE tests including dynamic driving conditions. Transportation Research Procedia 2019; 40: 338–345, https://doi.org/10.1016/j.trpro.2019.07.050.

7. Langbroek J H M, Cebecauer M, Malmsten J, Franklin J P, Susilo Y O, Georén P. Electric vehicle rental and electric vehicle adoption. Research in Transportation Economics 2019; 73: 72–82, https://doi.org/10.1016/j.retrec.2019.02.002.

8. Merkisz J, Pielecha J, Radzimirski S. New trends in emission control in the European Union. Springer Tracts on Transportation and Traffic 2014; 4: 170, https://doi.org/10.1007/978-3-319-02705-0.

9. Micari S, Polimeni A, Napoli G, Andaloro L, Antonucci V. Electric vehicle charging infrastructure planning in a road network. Renewable and Sustainable Energy Reviews 2017; 80: 98–108, https://doi.org/10.1016/j.rser.2017.05.022.

10. Muzi N. New car CO_2 standards: Is the job of securing electric cars in Europe done? Transport & Environment 2019. www.transportenvironment.org.

11. Pavlovic J, Marotta A, Ciuffo B. CO_2 emissions and energy demands of vehicles tested under the NEDC and the new WLTP type approval test procedures. Applied Energy 2016; 177: 661–670, https://doi.org/10.1016/j.apenergy.2016.05.110.

12. Pielecha I, Cieslik W, Szalek A. Operation of electric hybrid drive systems in varied driving conditions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (1): 16–23, https://doi.org/10.17531/ein.2018.1.3.

13. Pielecha I, Cieslik W, Szalek A. Operation of hybrid propulsion systems in conditions of increased supply voltage. International Journal of Precision Engineering and Manufacturing 2017; 18: 1633–1639, https://doi.org/10.1007/s12541-017-0192-3.

14. PSPA, 2019. Licznik elektromobilności. Polskie Stowarzyszenie Paliw Alternatywnych. pspa.com.pl

15. Sun B, Zhang T, Ge W, Tan C, Gao S. Driving energy management of front-and-rear-motor-drive electric vehicle based on hybrid radial basis function. Archives of Transport 2019; 49 (1): 47–58, https://doi.org/10.5604/01.3001.0013.2775.

16. Tsokolis D, Tsiakmakis S, Dimaratos A, Fontaras G, Pistikopoulos P, Ciuffo B, Samaras Z. Fuel consumption and CO_2 emissions of passenger cars over the New Worldwide Harmonized Test Protocol. Applied Energy 2016; 179: 1152–1165, https://doi.org/10.1016/j.apenergy.2016.07.091.

17. Wei Z, Xu Z, Halim D. Study of HEV power management control strategy based on driving pattern recognition. Energy Procedia 2016; 88: 847–853.

https://doi.org/10.1016/j.egypro.2016.06.062.

18. Wu W, Freese D, Cabrera A, Kitch W A. Electric vehicles' energy consumption measurement and estimation. Transportation Research Part D: Transport and Environment 2015; 34: 52–67, https://doi.org/10.1016/j.trd.2014.10.007.

19. Xie L, Luo Y, Zhang D, Chen R, Li K. Intelligent energy-saving control strategy for electric vehicle based on preceding vehicle movement. Mechanical Systems and Signal Processing 2019; 130: 484–501. https://doi.org/10.1016/j.ymssp.2019.05.027.

20. Zhang S, Gajpal Y, Appadoo S S, Abdulkader M M S. Electric vehicle routing problem with recharging stations for minimizing energy consumption. International Journal of Production Economics 2018; 203: 404–413, https://doi.org/10.1016/j.ijpe.2018.07.016.

mgr inż. Marta Mielczarek[1]

dr hab. inż. Mieczysław Słowik, prof. PP[1]

dr hab. Karol Andrzejczak[2]

[1] Institute of Civil Engineering
Poznan University of Technology, ,
Piotrowo St. 5, 60-965 Poznań, Poland
E-mail: marta.mielczarek@put.poznan.pl
mieczyslaw.slowik@put.poznan.pl

[2] Institute of Mathematics
Poznan University of Technology, ,
Piotrowo St. 3A, 60-965 Poznań, Poland
E-mail: karol.andrzejczak@put.poznan.pl

Ocena wpływu zawartości elastomeru styren-butadien-styren na właściwości funkcjonalne lepiszczy asfaltowych

Keywords: *dynamic shear rheometer (DSR), rutting factor, copolymer SBS, bitumen, complex shear modulus*

Słowa kluczowe: *reometr dynamicznego ścinania, wskaźnik odkształcalności, kopolimer SBS, asfalt, dynamiczny moduł ścinania*

Abstract : Tematyka pracy związana jest z zagadnieniem polepszenia właściwości funkcjonalnych drogowych nawierzchni asfaltowych poprzez modyfikację lepiszcza asfaltowego kopolimerem SBS. Głównym celem pracy jest ocena odporności na odkształcenia trwałe oraz wrażliwości na zmiany temperatury asfaltów drogowych modyfikowanych polimerami, które są najczęściej używane w wierzchnich warstwach konstrukcji nawierzchni drogowych i lotniskowych. Przedmiotem badań były asfalty pochodzące z różnych złóż ropy naftowej (rosyjskiej i wenezuelskiej). Asfalty te poddano modyfikacji w warunkach laboratoryjnych z dodatkiem koncentratu o znanej zawartości kopolimeru SBS równej 9%. Otrzymano w ten sposób lepiszcza asfaltowe o znanej zawartości kopolimeru SBS równej 1,5%; 3,0%; 4,5% oraz 6%. Właściwości reologiczne badanych asfaltów oznaczono z użyciem reometru dynamicznego ścinania DSR stosując w testach obciążenie sinusoidalnie zmienne, w szerokim zakresie temperatury pomiarowej (od 40°C do 100°C). Analizując wartości dynamicznego modułu ścinania |G*| wszystkich badanych asfaltów można stwierdzić, iż wzrost zawartości kopolimeru SBS w badanym lepiszczu zwiększa wartość |G*|, co może skutkować większą odpornością na odkształcenia trwałe nawierzchni drogowej spowodowane wielokrotnie powtarzającymi się obciążeniami ruchem pojazdów, w szczególności w przypadku nawierzchni eksploatowanej w wysokiej temperaturze Odporność mieszanek mineralno-asfaltowych (MMA) na powstawanie kolein jest jednym z podstawowych parametrów związanych z eksploatacją nawierzchni drogowych, wpływając zarówno na bezpieczeństwo, jak i komfort jazdy użytkowników.

1. Wprowadzenie

Nawierzchnie drogowe w dzisiejszych czasach są poddawane coraz to większym obciążeniom wywołanym ruchem drogowym [4]. Zwracając uwagę na koszty zarówno podczas budowy, jak i eksploatacji drogi [6], należałoby optymalizować m.in. skład materiałów, z których wykonana jest konstrukcja nawierzchni drogowej. Analizując wyniki badań przedstawione w pracach [1,11,12] można stwierdzić, iż jednym z kluczowych czynników, wpływających na powstawanie kolein w nawierzchniach drogowych jest skład mieszanki mineralno-asfaltowej (MMA), a zwłaszcza zastosowane lepiszcze asfaltowe. W związku z tym dąży się, aby uzyskać jak najlepsze właściwości reologiczne stosowanych asfaltów uzyskiwanych w procesie destylacji ropy naftowej. Poprawę tych właściwości uzyskuje się poprzez wprowadzenie do struktury lepiszcza różnego rodzaju modyfikatorów, tj. polimery [5], rozdrobniona guma z opon samochodowych [6], czy też asfalty naturalne [7]. W licznych artykułach naukowych analizie poddano efekty modyfikacji asfaltu najczęściej używanymi polimerami, wśród nich plastomery (np. polietylen, polipropylen, etylen octan winylu [17]), elastomery termoplastyczne: SBS (styren-butadien-styren)) [1,16]; SIS (styren izopren styren) lub modyfikatory mieszane składające się z kilku polimerów [2,12]. Lepiszcza modyfikowane polimerami wykazują poprawę właściwości reologicznych w porównaniu z asfaltami niemodyfikowanymi [20]. Najczęściej używany w budownictwie drogowym jest kopolimer blokowy SBS, po dodaniu do gorącego asfaltu zwiększa swoją objętość kilkukrotnie w stosunku do wartości początkowej [21]. W przypadku gdy stężenie polimeru w asfalcie modyfikowanym wynosi około 6%, polimer ten stanowi fazę rozpraszającą i tworzy ciągłą sieć w strukturze asfaltu. Przy mniejszym stężeniu elastomeru SBS, sieć polimeru może nie być ciągła. Dlatego bardzo ważne jest, zarówno ze względów technicznych jak i ekonomicznych, określenie granicznej zawartości kopolimeru SBS w lepiszczu asfaltowym, przy którym tworzy się ciągła sieć polimeru. Elastomeroasfalty w temperaturach eksploatacyjnych charakteryzują się sprężystością natychmiastową (odkształcenie sprężyste) i opóźnioną (pełzanie) [19,21]. W pracach [1,12,21] przedstawiono sposoby modyfikacji oraz korzyści jakie niesie zastosowanie kopolimeru SBS do modyfikacji lepiszcza, tj. przyrost temperatury mięknienia, zmniejszenie wrażliwości temperaturowej (następuje rozszerzenie temperaturowego zakresu lepkosprężystości), zwiększenie kohezji w niskiej temperaturze, powoduje znaczne polepszenie właściwości sprężystych (obserwowane np. w teście nawrotu sprężystego). Poprawa właściwości reologicznych lepiszczy przekłada się na właściwości uzyskanych MMA, w których zastosowano lepiszcze modyfikowane. [13] tj. zwiększenie odporności na odkształcenia trwałe i pękanie indukowane termicznie.

Airey w pracy [1] analizie poddał asfalty pochodzące z dwóch złóż ropy naftowej (rosyjskiej i wenezuelskiej). Do modyfikacji asfaltów wykorzystał kopolimer SBS, uzyskując stężenia polimeru w asfalcie odpowiednio 3%, 5% i 7%. Zauważył istotny wpływ polimeru na właściwości reologiczne asfaltu modyfikowanego; tj. przyrost wartości dynamicznego modułu ścinania oraz większy udział części sprężystej asfaltu zwłaszcza w wysokiej temperaturze. Ukazał również problem kompatybilności układu asfalt-polimer. Airey [1] wykazał, że asfalty parafinowe (pochodzenia rosyjskiego) poprzez większą zawartość części aromatycznych lepiej wiążą polimer w strukturze asfaltu modyfikowanego. Behnood i Olek [6] dokonali analizy porównawczej trzech rodzajów modyfikatorów: kopolimeru SBS, gumy oraz kwasu polifosforowego. Do badań właściwości niskotemperaturowych wykorzystali reometr zginanej belki (BBR), natomiast właściwości reologiczne w wysokich temperaturach określili przy użyciu reometru DSR. W wysokich temperaturach zaobserwowali przyrost wartości dynamicznego modułu ścinania dla wszystkich badanych asfaltów. Asfalty wykorzystywane do produkcji MMA, w nawierzchniach drogowych są narażone na procesy starzenia zarówno podczas magazynowania, produkcji mieszanki mineralno-asfaltowej,

transportu, wbudowania, jak również podczas eksploatacji nawierzchni [22]. Jako najbardziej niekorzystne ze względu na wysokie temperatury uznaje się starzenie występujące podczas produkcji i wbudowania MMA [24]. Podczas starzenia krótkoterminowego (technologicznego) zachodzącego w czasie procesu produkcji i budowy nawierzchni asfaltowej, asfalt poddawany jest działaniu wysokiej temperatury (140 - 200˚C) oraz tlenu zawartego w powietrzu. Airey [1] oraz Sarnowski [19] ukazali problem związany ze starzeniem asfaltów modyfikowanych. Lepiszcza modyfikowane wykazują poprawę właściwości reologicznych w szerokim zakresie lepkosprężystym. Autorzy zaobserwowali, iż po starzeniu asfaltów modyfikowanych występuje większy udział części lepkiej w stosunku do części sprężystej, co może być spowodowane częściową degradacją polimeru w wysokiej temperaturze, jaka występuje podczas procesów technologicznych otaczania kruszywa asfaltem, transportu, wbudowania i zagęszczania MMA [1]. Następny etap starzenia występuje podczas eksploatacji nawierzchni, jest to tzw. starzenie długoterminowe (eksploatacyjne). Lepiszcze narażone jest w tym przypadku na temperatury w okresie letnim dochodzące do 60˚C, przy jednoczesnym oddziaływaniu tlenu, promieniowania słonecznego, wody oraz środków chemicznych [16,22]. Bai [6] wykonał badania wpływu starzenia krótkoterminowego i długoterminowego na właściwości reologiczne asfaltu modyfikowanego kopolimerem SBS. Badaniom poddano trzy stężenia polimeru w asfalcie, tj. 3%, 6% oraz 9%. Na podstawie badań przy użyciu reometru dynamicznego ścinania DSR, aparatu Fraassa oraz penetrometru wykazali negatywny wpływ starzenia na właściwości niskotemperaturowe.

Potrzeby w zakresie utrzymania oraz remontów sieci dróg są bardzo duże. Systematycznie zwiększające się obciążenie ruchem wpływa bardzo niekorzystnie na stan nawierzchni przyśpieszając ich degradację [15]. Duże znaczenie w poznaniu stanu dróg odgrywa diagnostyka nawierzchni i badania cech funkcjonalnych tj. równość podłużna, równość poprzeczna (koleiny) [8], współczynnik tarcia, nośność itp. [14,15,23]. Do podstawowych rodzajów zniszczenia nawierzchni asfaltowych można zaliczyć: koleinowanie, pękanie zmęczeniowe oraz pękanie niskotemperaturowe [10,13]. Na powstawanie odkształceń trwałych (kolein) w nawierzchniach drogowych wpływ ma wiele czynników [13,14] m.in. zastosowane kruszywo, lepiszcze, mieszanka mineralno-asfaltowa (MMA), występujące warunki klimatyczne, obciążenie ruchem [9] czy też zastosowana konstrukcja nawierzchni.

Celem niniejszego artykułu jest analiza właściwości funkcjonalnych lepiszczy modyfikowanych kopolimerem SBS ze szczególnym uwzględnieniem wrażliwości temperaturowej, gdyż rodzaj zastosowanego lepiszcza jest jednym z kluczowych czynników mających wpływ na odporność na powstawanie odkształceń trwałych (kolein) w nawierzchniach asfaltowych. Odporność MMA na powstawanie kolein jest jednym z podstawowych warunków prawidłowej eksploatacji nawierzchni drogowych, wpływając zarówno na bezpieczeństwo, jak i komfort jazdy. Oryginalnym osiągnięciem niniejszej pracy było zastosowanie analizy wrażliwości temperaturowej badanych lepiszczy modyfikowanych w szerokim zakresie temperatur.

2. Charakterystyka badanych lepiszczy

Modyfikacja asfaltów kopolimerem SBS z reguły odbywa się w rafineriach, rzadziej zaś w instalacjach koncernów drogowych. Asfalt modyfikowany polimerem wykorzystywany do produkcji mieszanki mineralno-asfaltowej można uzyskać poprzez zakup z rafinerii gotowego lepiszcza modyfikowanego, produkcję asfaltu modyfikowanego w specjalnej instalacji technologicznej lub zakup asfaltu o znanej zawartości masowej kopolimeru SBS np. 9% i wymieszanie go w odpowiednich proporcjach z ponaftowym asfaltem, drogowym [21].

Badania wykonano wykorzystując asfalty 50/70 o zbliżonej twardości, wyrażonej poprzez wartość penetracji oznaczonej w 25˚C (tab.1), wyprodukowane z ropy naftowej pochodzącej z Wenezueli oraz Rosji. W przeprowadzonych badaniach asfalty połączono z koncentratem asfaltu modyfikowanego kopolimerem SBS (kopolimer blokowy o strukturze liniowej) o zawartości polimeru 9%, poprzez komponowanie w proporcjach 5:1, 2:1, 1:1 oraz 1:2 otrzymując odpowiednio asfalt o zawartości kopolimeru SBS: 1,5%; 3,0%; 4,5% i 6,0% (w stosunku do masy otrzymanego asfaltu modyfikowanego). Badane lepiszcza asfaltowe oznakowano w pracy poprzez podanie pochodzenia asfaltu, a następnie zawartości procentowej kopolimeru SBS, np.:

- R6%SBS – oznacza asfalt wyprodukowany z rosyjskiej ropy naftowej o zawartości 6,0% kopolimeru SBS,
- V50/70 – oznacza asfalt 50/70 pochodzący z wenezuelskiej ropy naftowej, nie zawierający kopolimeru SBS,
- K9%SBS – oznacza koncentrat asfaltu modyfikowanego zawierający 9,0% SBS.

Analizie zostały poddane asfalty zarówno w stanie wyjściowym, jak i po procesie starzenia technologicznego (krótkoterminowego), symulowanego metodą RTFOT (Rolling Thin Film Oven Test) wg PN-EN 12607-1:2014.

Tabela 1. Podstawowe właściwości badanych lepiszczy asfaltowych

Właściwości / Badany materiał	T_{PiK} [°C]	Pen_{25} [mm/10]
V50/70	47,4±0,2	66,0±0,8
V1,5%SBS	47,9±0,3	69,5±0,4
V3%SBS	52,8±1,3	71,3±0,8
V4,5%SBS	74,5±2,0	66,4±0,6
V6%SBS	87,9±1,3	66,3±0,8
R50/70	47,8±0,4	69,3±0,3
R1,5%SBS	48,8±0,3	70,3±0,3
R3%SBS	49,5±0,3	71,3±0,4
R4,5%SBS	77,0±1,1	69,5±0,5
R6%SBS	83,8±06	71,9±1,0
K9%SBS	100,3±1,4	74,3±1,0

gdzie: T_{PiK} -temperatura mięknienia wg PN-EN 1427:2015-08,
Pen_{25} - penetracja w 25°C wg PN-EN1426:2015-08

Analizując wyniki przedstawione w tabeli nr 1 lepiszcza asfaltowe były dobrane w taki sposób, aby uzyskać asfalty o zbliżonej twardości, wyrażonej poprzez penetrację w 25°C (Pen_{25} uzyskano w przedziale 66,0 mm/10 do 74,3 mm/10). Wszystkie badane asfalty modyfikowane można więc zakwalifikować do klasy asfaltów modyfikowanych 45/80 dostępnych na polskim rynku, mimo iż posiadają różne procentowe zawartości kopolimeru SBS.

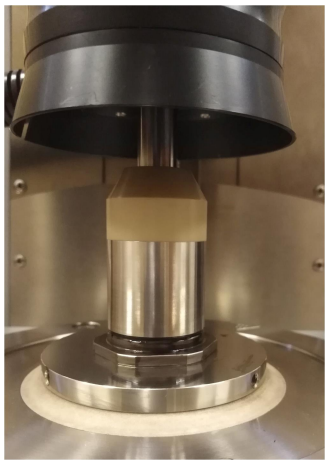
3. Cel i metodyka badań

Głównym celem pracy jest ocena odporności na odkształcenia trwałe oraz wrażliwości temperaturowej asfaltów drogowych modyfikowanych kopolimerem SBS na podstawie badań przeprowadzonych przy pomocy reometru dynamicznego ścinania (DSR) typu Physica MCR 101 (rys. 2). Odkształcenia trwałe oraz wrażliwość na zmiany temperatury w strefie klimatycznej obejmującej Polskę mają kluczowe znaczenie w eksploatacji nawierzchni asfaltowych.

Badania przeprowadzono zgodnie z normą PN-EN 14770:2012 „Asfalty i lepiszcza asfaltowe. Oznaczenie zespolonego modułu ścinania i kąta przesunięcia fazowego w reometrze dynamicznego ścinania (DSR)" dwiema metodami, w których zastosowano wymuszenie kinematyczne (sinusoidalne):

a) przy różnych zakresach częstości kątowej od 100 rad/s do 0,1 rad/s oraz stałej temperaturze badania równej 60°C±0,01˚C,
b) o amplitudzie kąta wychylenia wrzeciona równej 10 mrad oraz przy zmiennej temperaturze, tj. od 100°C do 40°C, przy czym co 1 min następowało obniżenie temperatury o 1°C. W tej procedurze badawczej przyjęto stałą wartość częstości kątowej równą 10 rad/s.

Próbkę lepiszcza asfaltowego umieszczano pomiędzy dwiema okrągłymi płytami równoległymi o średnicy Ø25mm, przy zachowaniu zadanej wysokości szczeliny równej 1 mm (rys. 1)



Rys. 1. Widok z boku próbki badanego asfaltu

Rys. 2. Reometr dynamicznego ścinania DSR typu Physica MCR 101

Zgodnie z założeniami amerykańskiej specyfikacji Superpave, podatność lepiszczy asfaltowych na powstawanie odkształceń trwałych w nawierzchniach drogowych określa się za pomocą współczynnika odkształcalności (rutting factor), wyrażonego poprzez wielkość |G*|/sin δ. W pracy wyznaczono ww. współczynnik dla asfaltów zarówno nie poddanych, jak i poddanych starzeniu krótkoterminowemu metodą RTFOT.

Analizie poddano również wartości indeksu modułu ścinania (SMI – Shear Modulus Index), będącego miarą wrażliwości temperaturowej badanych lepiszczy, który obliczono wg wzoru [20]:

$$SMI_{T_2/T_1} = \left| \frac{\log\log\left|G^*_{T_1}\right| - \log\log\left|G^*_{T_2}\right|}{\log(T_1 + 273{,}15) - \log(T_2 + 273{,}15)} \right| \tag{1}$$

gdzie:

SMI - Shear Modulus Index

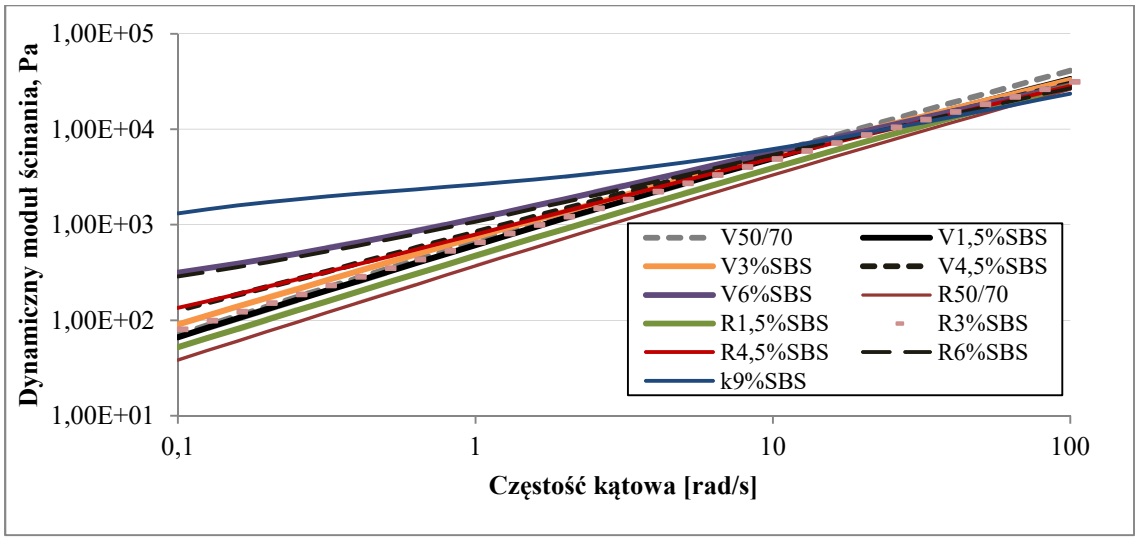
$\left|G^*_{T_1}\right|; \left|G^*_{T_2}\right|$ - dynamiczny moduł ścinania w temperaturze T_1, T_2, [Pa]

T_1; T_2 – ekstremalne temperatury pomiarów wykonanych w reometrze dynamicznego ścinania DSR, przy czym $T_1 > T_2$, [°C]

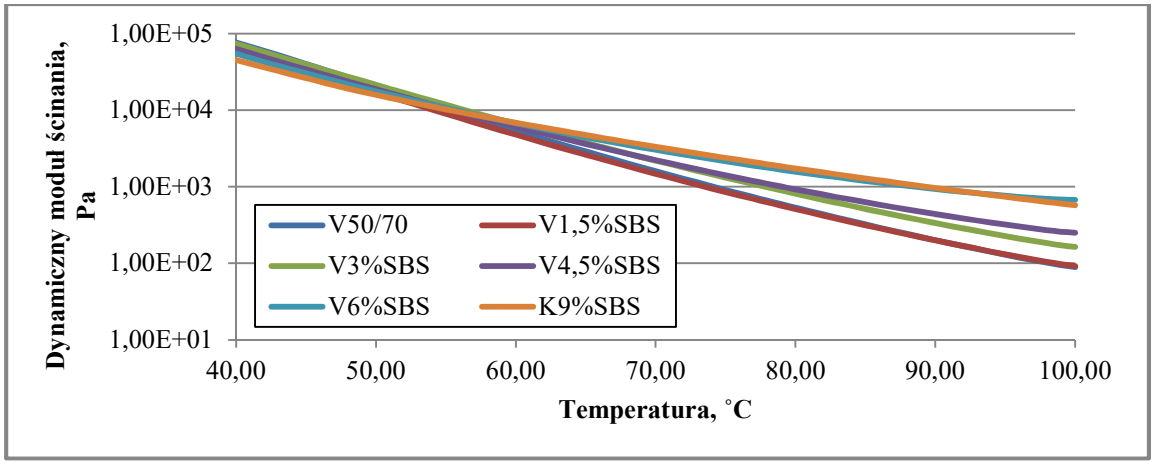
W niniejszej pracy przyjęto $T_1 = 100°C$; $T_2 = 40°C$.

4. Analiza wyników badań

Rysunek nr 3 przedstawia wykres zależności dynamicznego modułu ścinania badanych asfaltów od częstości kątowej (zakres od 0,1 rad/s do 100 rad/s). Dynamiczny moduł ścinania wzrasta wraz ze wzrostem częstości kątowej dla wszystkich analizowanych asfaltów. Największą wartość |G*| przy częstości kątowej 0,1 rad/s posiada asfalt o zawartości kopolimeru SBS równej 9%, natomiast najmniejszą asfalt R50/70. Przy częstości kątowej równej 100 rad/s dynamiczny moduł ścinania uzyskuje wartości na zbliżonym poziomie od wartości równej 23680Pa dla K9%SBS do 41010Pa dla asfaltu V50/70. Reasumując wzrost zawartości kopolimeru SBS w lepiszczu asfaltowym powoduje przyrost wartości dynamicznego modułu ścinania, przy częstości kątowej 10 rad/s.

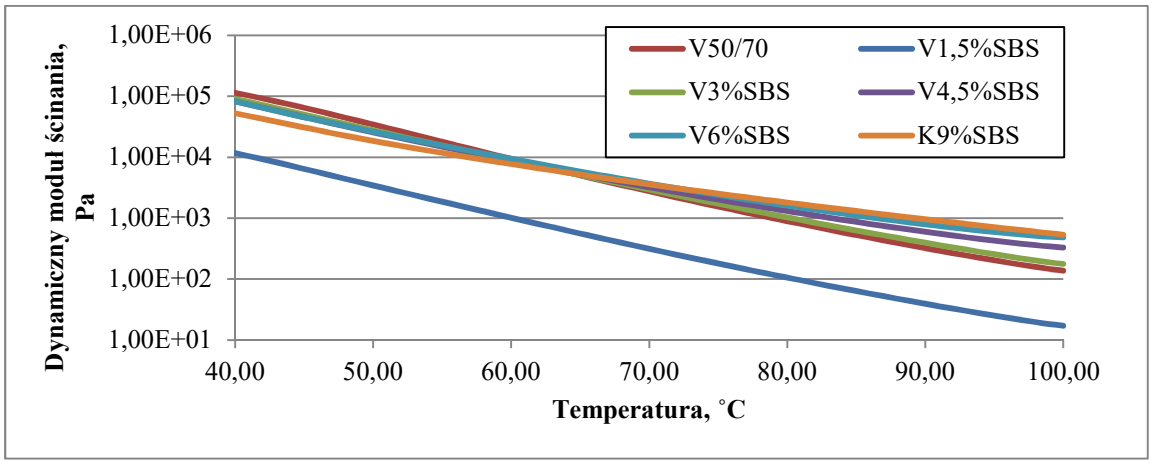


Rys. 3. Wykres zależności dynamicznego modułu ścinania od częstości kątowej asfaltów badanych w 60˚C

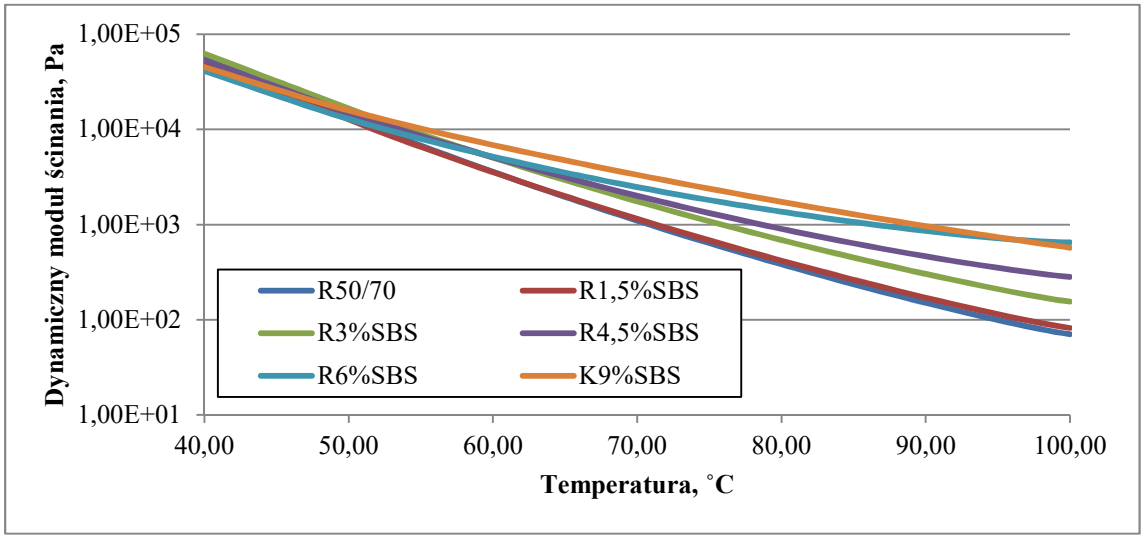


Rys. 4. Wykres zależności dynamicznego modułu ścinania |G*| od temperatury dla asfaltów pochodzenia wenezuelskiego niepoddanych starzeniu, przy stałej częstości kątowej równej 10

rad/s

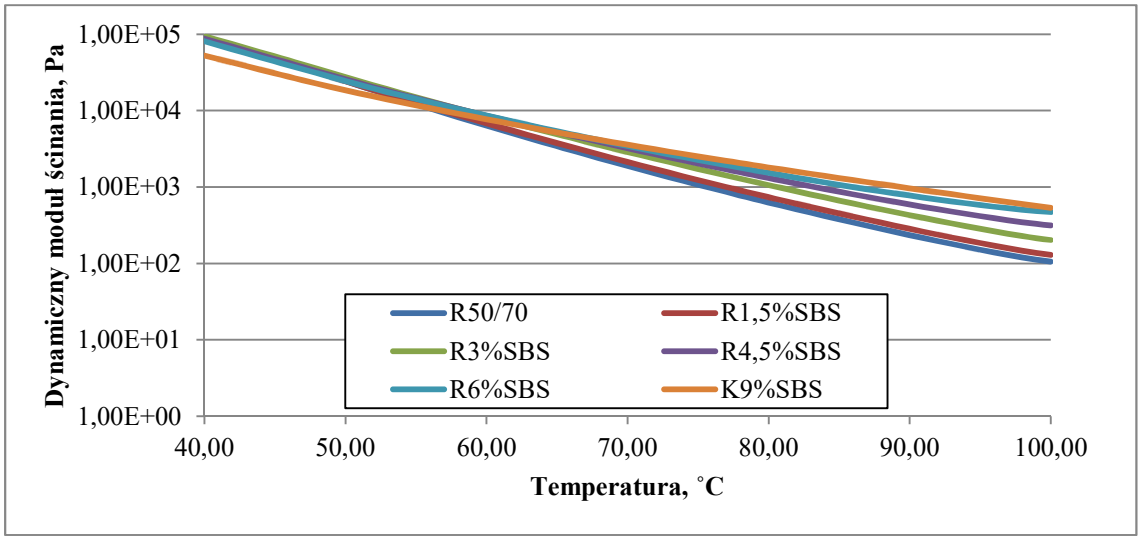


Rys.5. Wykres zależności dynamicznego modułu ścinania |G*| od temperatury dla asfaltów pochodzenia wenezuelskiego poddanych starzeniu RTFOT, przy stałej częstości kątowej równej 10rad/s



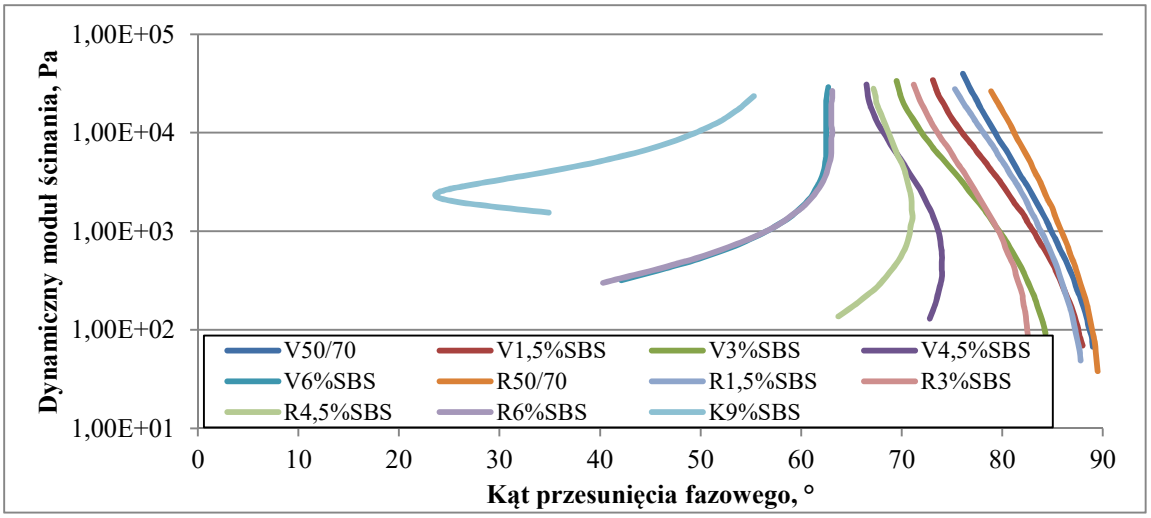
Rys.6. Wykres zależności dynamicznego modułu ścinania |G*| od temperatury dla asfaltów pochodzenia rosyjskiego niepoddanych starzeniu, przy stałej częstości kątowej równej10 rad/s

Lepiszcza asfaltowe są materiałami o właściwościach lepkosprężystych. Analizując wartości kąta przesunięcia fazowego δ można ocenić zmiany właściwości reologicznych asfaltu w całym spektrum temperatur zarówno podczas procesu produkcji, wbudowania, jak i eksploatacji nawierzchni asfaltowej. Materiały lepkie charakteryzują się tym, że współczynnik tłumienia oznaczany jako $\mathrm{tg}\delta \rightarrow \infty$ ($\delta = 90°$), natomiast w przypadku materiałów sprężystych $\mathrm{tg}\delta = 0$ ($\delta = 0°$); materiały lepkosprężyste mają kąt przesunięcia fazowego o wartości zawierającej się w przedziale $0° < \delta < 90°$. Na rys. 4-7 przedstawiono wykresy zależności dynamicznego modułu ścinania |G*| od temperatury badanych lepiszczy asfaltowych pochodzenia wenezuelskiego oraz rosyjskiego, zarówno asfaltów niepoddanych starzeniu, jak i po procesie starzenia RTFOT. Wraz ze wzrostem temperatury badania maleje wartość dynamicznego modułu ścinania |G*| dla wszystkich analizowanych asfaltów.



Rys.7. Wykres zależności dynamicznego modułu ścinania |G*| od temperatury dla asfaltów pochodzenia rosyjskiego poddanych starzeniu RTFOT, przy stałej częstości kątowej równej 10rad/s

Ze względu na wysokie temperatury panujące w okresie letnim istotna jest składowa sprężysta, co wiąże się z małymi wartościami tgδ. Podczas badań zaobserwowano, iż w przypadku asfaltów niemodyfikowanych (R i V) oraz asfaltów o małej zawartości kopolimeru SBS (do 3%), występuje prawidłowość, że im większe wartości dynamicznego modułu ścinania, tym mniejsze wartości kąta przesunięcia fazowego δ, co obrazuje wykres Black'a (rys. 8). Wykresy przedstawiające zależność dynamicznego modułu ścinania od kąta przesunięcia fazowego, zwane wykresami Black'a, pozwalają na wykonanie analizy dwóch podstawowych parametrów wyznaczonych w reometrze DSR [1,3,20]. W przypadku asfaltów o zawartości kopolimeru SBS 6% i 9% uzyskuje się małe wartości kąta przesunięcia fazowego, zarówno przy bardzo małych, jak i przy dużych wartościach |G*|. Największą zmienność wartości kąta przesunięcia fazowego zaobserwowano dla asfaltów referencyjnych 50/70 oraz asfaltów niskomodyfikowanych (o zawartości kopolimeru do 3%). Wartości te w wysokich temperaturach są bliskie 90° można więc uznać, że lepiszcza te w zakresie wysokich temperatur mają właściwości zbliżone do cieczy lepkiej. Zwiększenie zawartości kopolimeru SBS w asfalcie powoduje, że zróżnicowanie wartości δ jest coraz mniejsze. Powyżej temperatury 70°C następuje zmniejszenie wartości δ dla asfaltów o zawartości kopolimeru 4,5%; 6% oraz dla koncentratu 9% SBS, co ukazuje korzystny wpływ zastosowania do modyfikacji polimeru, ponieważ lepiszcza modyfikowane posiadają w wysokich temperaturach większy udział części sprężystej co może świadczyć o większej odporności na odkształcenia trwałe.



Rys.8. Wykres Black'a przedstawiający zależność dynamicznego modułu ścinania od kąta przesunięcia fazowego badanych lepiszczy asfaltowych

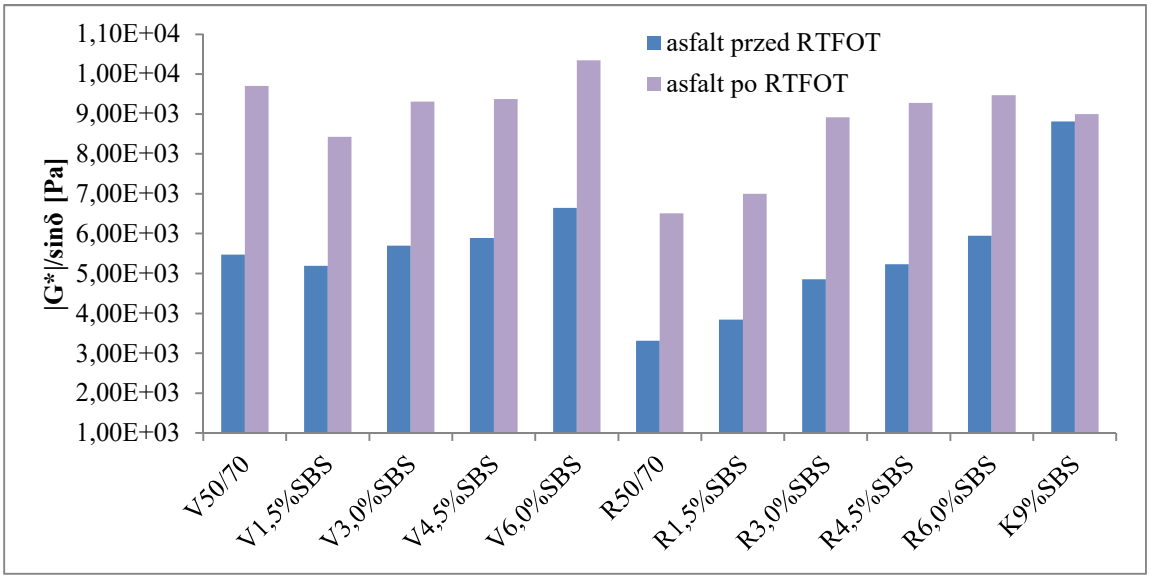
W badaniach wyznaczono wskaźnik odkształcalności (definiowany jako stosunek dynamicznego modułu ścinania |G*| do sinusa kąta przesunięcia fazowego δ (|G*|/sinδ)) w temperaturze 60°C, przyjętej jako temperatura ekstremalna występująca w nawierzchniach asfaltowych w Polsce (w której wykonuje się również test koleinowania MMA wg PN-EN 12697-22:2008).

Specyfikacja Superpave wskazuje na związek pomiędzy odpornością na powstawanie odkształceń trwałych w nawierzchniach asfaltowych, a właściwościami badanych lepiszczy oznaczonych w reometrze DSR wprowadzając następujące wymagania:

|G*|/sinδ ≥ 1,0 kPa - dla asfaltu niepoddanego starzeniu

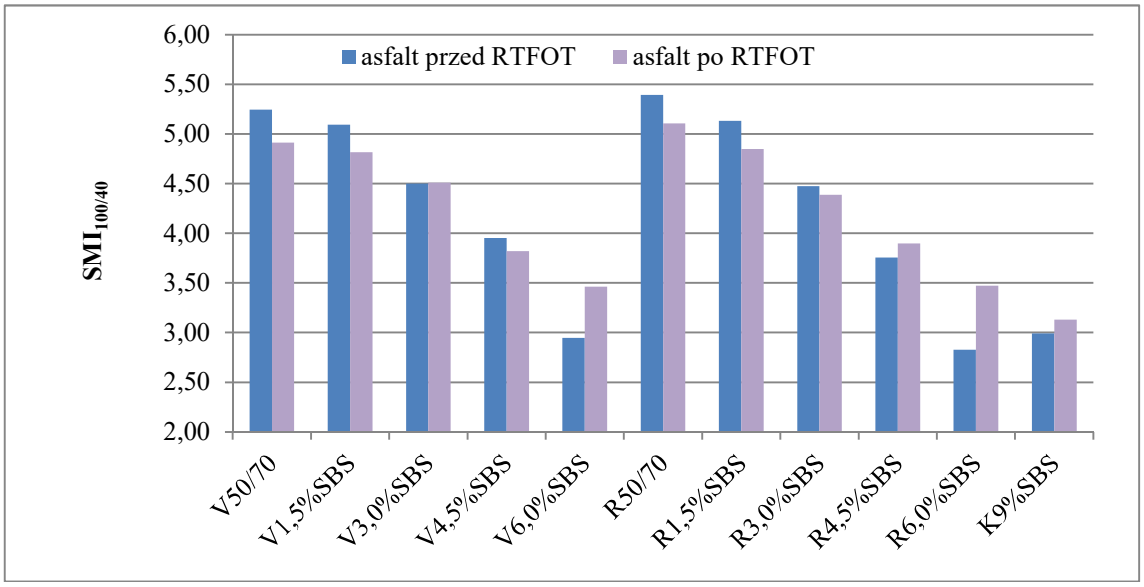
|G*|/sinδ ≥ 2,2 kPa - dla asfaltu poddanego starzeniu technologicznemu symulowanemu metodą RTFOT.

Analizując wyniki zamieszczone na rys.9, można uznać że wszystkie badane lepiszcza asfaltowe spełniają powyższe wymagania stawiane przez specyfikacje Superpave. Większa wartość wskaźnika odkształcalności asfaltu charakteryzującego odporność na odkształcenia trwałe nawierzchni asfaltowych uzyskuje się poprzez większą wartość dynamicznego modułu ścinania |G*| i mniejszą wartość kąta przesunięcia fazowego δ.



Rys.9. G*/sinδ badanych lepiszczy asfaltowych w temperaturze 60˚C.

Najmniejszą wartość wskaźnika odkształcalności zaobserwowano dla asfaltów niemodyfikowanych pochodzenia rosyjskiego zarówno przed jak i po starzeniu metodą RTFOT. Szczególną uwagę należy zwrócić na wartość |G*|/sinδ koncentratu o zawartości kopolimeru SBS 9%, ponieważ różnica wskaźnika odkształcalności przed i po starzeniu wynosi zaledwie 0,2kPa; co może świadczyć o małym wpływie starzenia na wartości wskaźnika odkształcalności (rutting factor).



Rys.10. Shear Modulus Index badanych lepiszczy asfaltowych

Miarą wrażliwości temperaturowej jest indeks penetracji PI. Może być wyznaczony na podstawie wyników penetracji w dwóch temperaturach lub metodą pośrednią wykorzystując penetrację w 25°C i temperaturę mięknienia. Powyższe metody pozwalają oszacować indeks penetracji asfaltów niemodyfikowancych. Natomiast w przypadku asfaltów modyfikowanych elastomerami, wyniki uzyskane z zastosowaniem każdej z metod mogą znacznie się różnić [20]. W niniejszej pracy wrażliwość temperaturową wyznaczono na podstawie wzoru (1). Można wnioskować, że założenia dotyczące penetracji w temperaturze mięknienia (800 mm/10) oraz temperaturze łamliwości (1,25 mm/10) dla asfaltów modyfikowanych elastomerami nie są słuszne [20]. Analiza uzyskanych wartości SMI (rys. 10) wykazała istotny wpływ zawartości kopolimeru na zmniejszenie wrażliwości asfaltów na zmiany sztywności w różnych temperaturach. Asfalt R50/70 okazał się najbardziej wrażliwy na zmiany właściwości pod wpływem zmiany temperatury. Najmniejszą wartość SMI w zakresie temperatur 100°C - 40°C osiągnął asfalt o zawartości kopolimeru SBS równej 6% w stanie wyjściowym (przed procesem starzenia symulowanego metodą RTFOT), natomiast po starzeniu wartość SMI wzrosła dla tej grupy asfaltów (zarówno pochodzenia wenezuelskiego, jak i rosyjskiego), co może świadczyć o częściowym rozpadzie polimeru pod wpływem wysokiej temperatury (163°C) i dopływu tlenu (starzenie metodą RTFOT).

5. Wnioski

Porównując wartości dynamicznego modułu ścinania badanych asfaltów można stwierdzić, iż wraz ze wzrostem zawartości kopolimeru SBS zwiększa się wartość |G*|, co może świadczyć o większej odporności nawierzchni asfaltowych wykonanych z użyciem

asfaltu modyfikowanego kopolimerem SBS na deformacje wywoływane powtarzającymi się naprężeniami ścinającymi (co obrazuje w warunkach rzeczywistych powtarzające się cykle obciążeń wywołane ruchem pojazdów).

Wraz ze wzrostem zawartości kopolimeru SBS w badanych asfaltach zmniejsza się wartość kąta przesunięcia fazowego, co skutkuje polepszeniem ich właściwości sprężystych.

Analiza wartości SMI (Shear Modulus Index) wykazała znaczący wpływ zawartości kopolimeru styren-butadien-styren (SBS) na zmniejszenie wrażliwości lepiszczy asfaltowych na zmiany sztywności przy zmiennej temperaturze. Zmniejszenie podatności na zmiany temperatury obrazuje bardzo korzystny wpływ zastosowania jako modyfikatora kopolimeru SBS w lepiszczu asfaltowym.

Zastosowanie w MMA lepiszczy modyfikowanych kopolimerem SBS poprawia właściwości funkcjonalne nawierzchni podatnych (co potwierdzają wartości wskaźnika odkształcalności oraz SMI), wpływając tym samym na poprawę parametrów eksploatacyjnych nawierzchni drogowych i tym samym ich trwałość.

References

1. Airey G. Rheological properties of styrene butadiene styrene polymer modified road bitumens. Fuel 2003; 82: 1709-1719.

2. Ahmedzade P. The investigation and comparison effects of SBS and SBS with new reactive thermopolymer on the rheological properties of bitumen. Construction and Building Materials 2013; 38: 285-291.

3. Andriescu A, Hesp S.A.M. Time-temperature superposition in rheology and ductile failure of asphalt binders. International Journal of Pavement Engineering 2009; 10(4): 229-240.

4. Andrzejczak K. Zmiany wzrostu wskaźnika nasycenia samochodami osobowymi. Wiadomości Statystyczne 2012; 11: 22-33.

5. Bai M. Investigation of low-temperature properties of recycling of aged SBS modified asphalt . Construction and Building Materials 2017; 150: 766-773

6. Behnood A, Olek J. Rheological properties of asphalt binders modifies with styrene-butadiene-styrene, ground tire rubber (GTR), or polyphosphoric acid (PPA). Construction and Building Materials 2017; 151: 464 - 478

7. Bilski M, Słowik M. Impact of aging on Gilsonite and Trinidad Epuré modified binders resistance to cracking. Bituminous Mixtures and Pavements VII 2019; I: 65-70. ISBN 978-1-351-06326-5

8. Bogdański B, Słowik M. Analiza porównawcza odporności na koleinowanie mieszanek mineralno-asfaltowych z uwzględnieniem kryteriów oceny wg metody francuskiej (LCPC) i brytyjskiej (BS). Międzynarodowa Konferencja Naukowo-Techniczna "Nowoczesne technologie w budownictwie drogowym" Poznań 2001: 300-309.

9. Cheng Z, Remenyte-Prescott R. Two probabilistic life-cycle maintenance models for the deteriorating pavement. Eksploatacja i Niezawodność - Maintenance and Reliability 2018; 20(3): 394-404.

10. D'Angelo J, Kluttz R, Dongré R, Stephens K, Zanzotto L. Revision of the Superpave High Temperature Binder Specification: The Multiple Stress Creep Recovery Test. Journal of the Association of Asphalt Paving Technologists 2007; 76: 123-157.

11. Gajewski M, Sybilski D, Bańkowski W. The influence of binder rheological properties on asphalt mixture permanent deformation. The Baltic Journal of Road and Bridge Engineering 2015; 10(1): 54-60.

12. Gajewski M, Sybilski D, Bańkowski W, Wróbel A, Mirski K. Ocena odporności na deformacje trwałe mieszanek mineralno-asfaltowych na podstawie zaproponowanego parametru funkcjonalnego lepiszcza. Część 1. Badania lepiszczy. Drogownictwo 2009; 9: 279-283.

13. Gajewski M, Wróbel A, Jemioło S, Sybilski D. Wpływ właściwości reologicznych lepiszcza na koleinowanie MMA. XIV International Conference Computer Systems Aided Science 2010: 857-866.

14. Judycki J, Jaskuła P. Diagnostyka i modernizacja konstrukcji nawierzchni drogowych. Diagnostyka, monitoring i modernizacja eksploatowanych obiektów budowlanych 2010; 56: 233-252.

15. Levulyté L, Žuraulis V, Sokolovskij E. The research of dynamic characteristics of vehicle driving over road roughness. Eksploatacja i Niezawodność - Maintenance and Reliability 2014; 16(4): 518-525.

16. Lu X, Isacsson U. Effect of ageing on bitumen chemistry and rheology. Construction and Building Materials 2002; 16: 15-22.

17. Merijs-Meri R, Abele A, Zicans J, Haritinovs V. Development of polyolefine elastomer modified bitumen and characterization of its rheological and structural properties. Bituminous Mixtures and Pavements VII 2019; I: 51- 57

18. Radziszewski P. Wpływ modyfikacji elastomerem SBS na właściwości reologiczne lepiszczy asfaltowych. Polimery 2008: 559-563.

19. Sarnowski M. Rheological properties of road bitumen binders modifies with SBS polymer and polyphophosphoric acid. Roads and Bridges - Drogi i Mosty 2015; 1: 47-65.

20. Słowik M. Thermorheological Properties Of Styrene-Butadiene-Styrene (SBS) Copolymer Modified Road Bitumen. Procedia Engineering 2017; 208:145-150.

21. Słowik M. Wybrane zagadnienia lepkosprężystości drogowych asfaltów modyfikowanych zawierających elastomer SBS. Rozprawy series Publishing House of Poznan University of Technology 2013; 508.

22. Słowik M, Adamczak P. Ocena wpływu starzenia krótkoterminowego na właściwości asfaltów drogowych modyfikowanych elastomerem SBS. Roads and Bridges - Drogi i mosty 2007; 1: 41-58

23. Surblys V, Žuraulis V, Sokolovskij E. Estimation of road roughness from data of on-vehicle mounted sensors. Eksploatacja i Niezawodność - Maintenance and Reliability 2017; 19(3): 396-374.

24. Zicans J, Ivanova T, Merijs-MEri R, Berzina R, Haritonovs V. Aging behavior of bitumen and elastomer modified bitumen. Bituminous Mixtures and Pavements VII 2019; I: 46-50