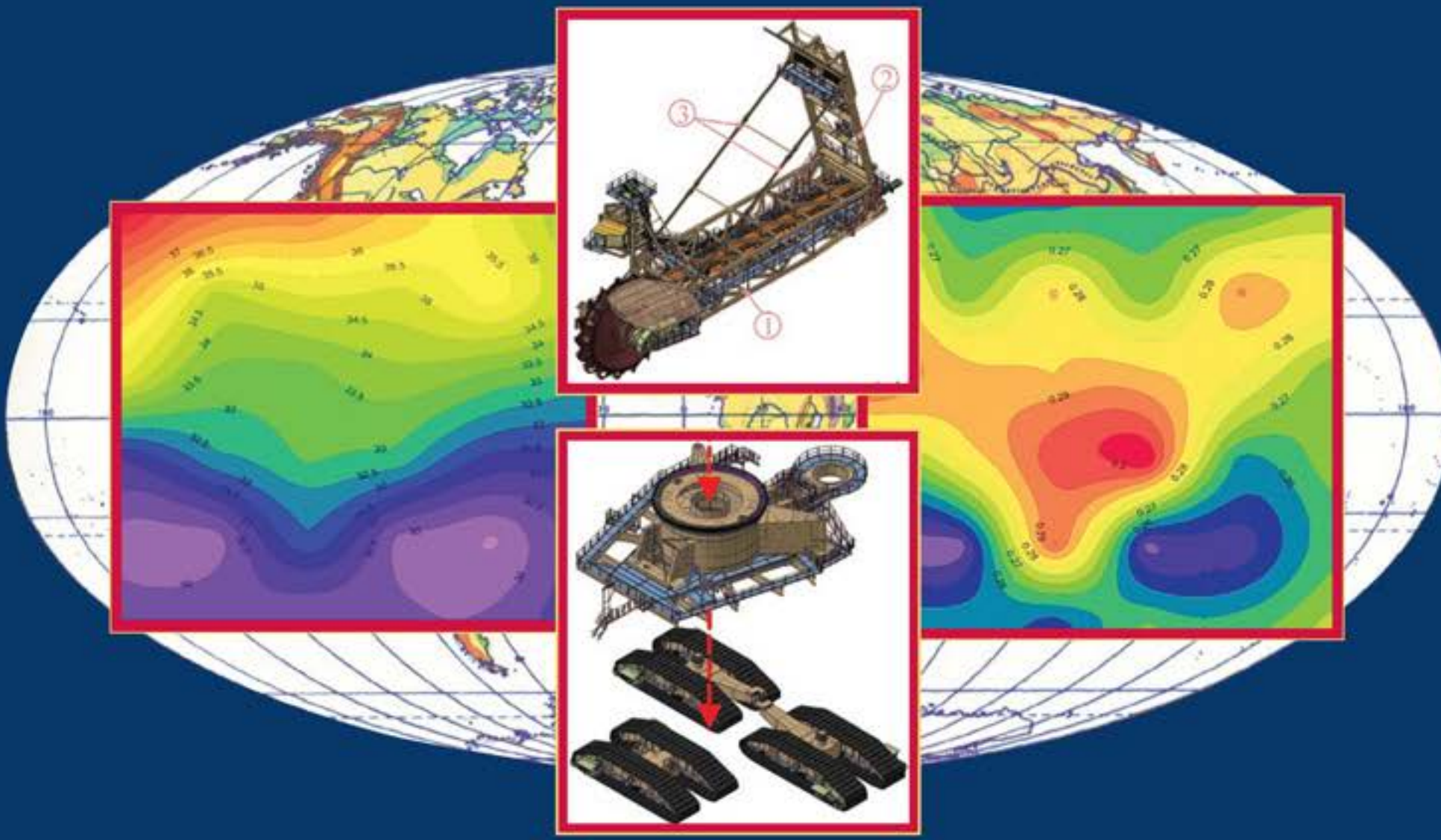
Vol. 20. No 2, 2018

ISSN 1507-2711
Cena: 25 zł (w tym 5% VAT)

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY



Polskie Naukowo Techniczne Towarzystwo Eksploatacyjne
Warszawa

Polish Maintenance Society
Warsaw

HAWRYLUK M, GRONOSTAJSKI Z, ZIEMBA J, DWORZAK Ł, JABŁOŃSKI P, RYCHLIK M. **Analysis of the influence of lubrication conditions on tool wear used in hot die forging processes**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 169–176, http://dx.doi.org/10.17531/ein.2018.2.01.

The study discusses the subject of lubrication in the processes of hot die forging with the consideration of the durability of forging tools and instrumentation. It presents a literature research as well as the authors' own investigations of the effect of the use of cooling and lubrication agents, the amount of the dosage as well as the direction of its application, and also the factors influencing the tribological conditions. The lubrication devices and systems currently applied in the industry have been analyzed as well. On this basis, making use of their knowledge and experience, the authors have developed and constructed a lubricating device. The elaborated system, implemented into the industrial process, makes it possible to select and ensure the optimal tribological conditions of the process by way of controlling the amount and frequency of the applied lubricant dosage. It can constitute an alternative for the manual method of lubricant application, which is dependent on the human factor, or the fully automated, yet expensive, lubrication systems. The obtained test results point to potential possibilities of a permanent introduction of the constructed device also into other forging processes, through its integration with a manipulator. The proposed solution ensures more stability and higher repeatability of the lubrication conditions as well as increases the efficiency of the production process, thus significantly reducing the unit costs of the production of forgings.

PUCHALSKI A, ŚLĘZAK M, KOMORSKA I, WIŚNIOWSKI P. **Multifractal analysis vehicle's in-use speed profile for application in driving cycles**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 177–181, http://dx.doi.org/10.17531/ein.2018.2.02.

Time signals recorded by the on-board diagnostic system (OBD), describing the manner of vehicle's movement in actual road conditions show non-stationarity and non-linearity, as well as statistical multiscalarity. In practice, it means that the analysis of registered time series requires modelling of non-linear phenomena. The aim of this study was to examine the nature of the vehicle speed profile in actual road conditions with the method of multifractal analysis. A number of studies indicates that the driving tests applied for many years have not been representative for the actual operating conditions of vehicles. For both the new Worldwide Harmonised Light duty Test Cycle (WLTC), a worldwide harmonised procedure of light vehicle testing, as well as in actual urban driving conditions along the measuring route, being subject to empirical research, confirmation of strong multifractal properties of the recorded vehicle speed time series have been obtained.

LUTY W. **Simulation-based analysis of the impact of vehicle mass on stopping distance**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 182–189, http://dx.doi.org/10.17531/ein.2018.2.03.

Results of experimental testing of motor truck tyres in dynamic braking conditions have been presented. With the measurement results being used as an example, higher normal wheel loads have been shown to result in longer time of rise in the longitudinal tangential tyre reaction force and in lower values of both the peak and sliding tyre-road adhesion coefficient. The data presented include results of simulation of the process of emergency braking of a motor truck whose mass can vary within wide limits. It can be seen from these results that an increase in the vehicle mass may considerably lengthen the vehicle stopping distance in emergency braking conditions.

Bošnjak S, Gnjatović N, Milenović I. **From 'a priori' to 'a posteriori' static stability of the slewing superstructure of a bucket wheel excavator**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 190–206, http://dx.doi.org/10.17531/ein.2018.2.04.

The complexity of the slewing superstructure (SS) balancing problem derives from the changeability of its geometric configuration, the complexity of working conditions as well as multiple limitations of the possible set of solutions. Having in mind the fact that the existing reference literature does not fully specify the procedure of static stability proof, the aforementioned procedure is presented in detail for the first time in this paper. A major contribution is represented in the classification of the slewing superstructure models into two groups which were named: the 'a priori' model (designed image of the SS) and the 'a posteriori' model (actual image of the SS). The fundamental stages of the 'a posteriori' model development method are presented in the paper. The transformation and validation of the calculation model 'a priori' to the calculation model 'a posteriori' was conducted on the basis of weighing results. A calculation of the basic parameters of the bucket wheel excavator (BWE) superstructure was conducted for both analyzed models by using the in-house developed software. The 'a posteriori' models provide a reliable calculation of the SS static stability and may be used not only for static stability proof of the machine as a whole, but also for load analysis of substructures and elements of BWE and related surface mining machines, such as spreaders. Besides that, the previously mentioned models are of extreme importance for a successful and reliable exploitation and maintenence of the machine since they present the basis for adjustment and control of limiting winch

HAWRYLUK M, GRONOSTAJSKI Z, ZIEMBA J, DWORZAK Ł, JABŁOŃSKI P, RYCHLIK M. **Analiza wpływu warunków smarowania na zużycie narzędzi w procesach kucia matrycowego na gorąco**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 169–176, http://dx.doi.org/10.17531/ein.2018.2.01.

Praca dotyczy problematyki smarowania w procesach kucia matrycowego na gorąco z uwzględnieniem trwałości narzędzi i oprzyrządowania kuźniczego. Przedstawiono badania literaturowe oraz własne autorów dotyczące wpływu zastosowania środków smarno-chłodzących, ilości dawki i kierunku jej podawania oraz innych czynników wpływających na warunki tribologiczne. Przeanalizowano także obecnie stosowane w przemyśle urządzenie i systemy smarowania. Na tej podstawie autorzy w oparciu o wiedzę i doświadczenie opracowali i zbudowali urządzenie smarujące. Opracowany system, zaimplementowany do przemysłowego procesu pozwala na dobór i zapewnienie optymalnych warunków tribologicznych w procesie poprzez sterowanie ilością i częstotliwością podawanej dawki środka smarnego. Może być ono alternatywą dla manualnej metody nanoszenia środka smarnego, zależnej od czynnika ludzkiego lub w pełni zautomatyzowanych, lecz drogich systemów smarowania. Uzyskane wyniki badań wskazują na potencjalne możliwości wprowadzenia na stałe do pracy zbudowanego urządzenia także do innych procesów kucia poprzez integrację z manipulatorem. Zaproponowane rozwiązanie zapewnia większą stabilność i powtarzalność warunków smarowania oraz pozytywnie wpływa na zwiększenie wydajności procesu wytwarzania, a tym samym znacząco obniża jednostkowe koszty produkcji odkuwek.

PUCHALSKI A, ŚLĘZAK M, KOMORSKA I, WIŚNIOWSKI P. **Multifraktalna analiza eksploatacyjnego profilu prędkości pojazdu w zastosowaniu do testów jezdnych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 177–181, http://dx.doi.org/10.17531/ein.2018.2.02.

Sygnały czasowe rejestrowane przez system diagnostyki pokładowej OBD i opisujące sposób ruchu pojazdu w rzeczywistych warunkach drogowych, wykazują niestacjonarność i nieliniowość oraz statystyczną wieloskalowość. W praktyce oznacza to, że analiza zarejestrowanych szeregów czasowych wymaga modelowania zjawisk nieliniowych. Celem niniejszej pracy było zbadanie charakteru profilu prędkości pojazdów w rzeczywistych warunkach drogowych metodą analizy multifraktalnej. Szereg badań wskazuje, że stosowane przez wiele lat testy jezdne nie były reprezentatywne dla rzeczywistych warunków eksploatacyjnych pojazdów. Zarówno dla nowego cyklu jezdnego WLTC, światowej zharmonizowanej procedury badań pojazdów lekkich jak i w rzeczywistych warunkach drogowych jazdy miejskiej na trasie pomiarowej, będącej przedmiotem badań doświadczalnych uzyskano potwierdzenie silnych własności multifraktalnych rejestrowanych szeregów czasowych prędkości pojazdu.

LUTY W. **Symulacyjna analiza wpływu masy pojazdu na drogę zatrzymania**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 182–189, http://dx.doi.org/10.17531/ein.2018.2.03.

W pracy przedstawiono wyniki badań eksperymentalnych ogumienia pojazdu ciężarowego w warunkach dynamicznego hamowania. Na przykładzie wyników pomiaru pokazano, że zwiększenie obciążenia normalnego koła skutkuje wzrostem czasu narastania wzdłużnej reakcji stycznej oraz spadkiem wartości współczynnika przyczepności opony do podłoża (przylgowej oraz poślizgowej). Przedstawiono wyniki symulacji procesu hamowania awaryjnego pojazdu ciężarowego, którego masa zmienia się znacząco. Wyniki wykazały ze zwiększenie masy pojazdu może istotnie wydłużyć jego drogę zatrzymania w warunkach hamowania awaryjnego.

Bošnjak S, Gnjatović N, Milenović I. **Równowaga statyczna nadwozia obrotowego koparki kołowej: od modelu "a priori" do modelu "a posteriori"**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 190–206, http://dx.doi.org/10.17531/ein.2018.2.04.

Złożoność problematyki stabilizacji nadwozia obrotowego (slewing superstructure, SS) wynika ze zmienności jego konfiguracji geometrycznej, złożoności warunków pracy oraz wielu ograniczeń możliwego zbioru rozwiązań. Ponieważ istniejąca literatura nie opisuje w pełni procedury przeprowadzania dowodu na równowagę statyczną, niniejsza praca stanowi pierwszą próbę opracowania takiej procedury. Głównym wkładem niniejszego artykułu jest klasyfikacja modeli nadwozia obrotowego na dwie grupy: model "a priori" (zaprojektowany obraz SS) i model "a posteriori" (rzeczywisty obraz SS). W artykule przedstawiono podstawowe etapy metody opracowywania modelu "a posteriori". Walidacji modelu obliczeniowego "a priori" i jego transformacji do modelu obliczeniowego "a posteriori" dokonano na podstawie wyników ważenia. Dla obydwu analizowanych modeli wykonano obliczenia podstawowych parametrów nadwozia koparki kołowej przy użyciu oprogramowania własnego. Modele "a posteriori" zapewniają niezawodne obliczenia równowagi statycznej SS i mogą być stosowane nie tylko do sprawdzania równowagi statycznej maszyny jako całości, ale również do analizy obciążenia podzespołów i elementów koparki kołowej oraz powiązanych maszyn górniczych, takich jak zwałowarki. Poza tym wspomniane wcześniej modele mają ogromne znaczenie dla skutecznej i niezawodnej eksploatacji i konserwacji maszyn, ponieważ stanowią podstawę do regulacji i kontroli granicznych wartości sił liny wciągarki, okresowej kontroli masy i

rope forces values, periodic control of mass and center of gravity position, as well as for a possible reconstruction which would be conducted in order to realize better customization of the machine versus operating conditions.

ZHU L, ZHANG Y, ZHANG R, ZHANG P. **Time-dependent reliability of spur gear system based on gradually wear process**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 207–218, http://dx.doi.org/10.17531/ein.2018.2.05.

To study dynamic evolution law of mechanical reliability caused by wear, gear transmission system is taken as a research object. Considering the effect of clearance caused by wear on gear teeth load in double meshing area, the formula of dynamic distribution load which is undertaken by two adjacent teeth is deduced. And the distributed pressure and meshing speed, which should be taken into account while calculating gear wear, are obtained based on the Winkler surface model and principle of tooth mesh. Based on the Archard's wear model, numerical simulation model for wear in spur gear is deduced, and the wear depth of each meshing points on teeth outline with different wear cycles are obtained. The calculation wear model is replaced with a surrogate model with Neural Network and Kriging method to overcome time-consuming defect. Random process model is integrated with the surrogate model, and dynamic reliability for nonlinear stochastic structure with unknown distribution characteristic is obtained with Neural Network-based Edgeworth series technique and four moment methods, which is compared with Kriging-based Monte Carlo simulation method. The computational efficiency and accuracy are also demonstrated.

ZHANG L, ZHANG J, ZHAI H, ZHOU S. **A new assessment method of mechanism reliability based on chance measure under fuzzy and random uncertainties**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 219–228, http://dx.doi.org/10.17531/ein.2018.2.06.

The traditional reliability analysis methods based on probability theory and fuzzy set theory have been widely used in engineering practice. However, these methods are unable directly measure the uncertainty of mechanism reliability with uncertain variables, i.e., subjective random and fuzzy variables. In order to address this problem, a new quantification method for the mechanism reliability based on chance theory is presented to simultaneously satisfy the duality of randomness and the subadditivity of fuzziness in the reliability problem. Considering the fact that systems usually have multilevel performance and the components have multimode failures, this paper proposes a chance theory based multi-state performance reliability model. In the proposed method, the chance measure is adopted instead of probability and possibility measures to quantify the mechanism reliability for the subjective probability or fuzzy variables. The hybrid variables are utilized to represent the random and fuzzy parameters, based on which solutions are derived to analyze the chance theory based mechanism reliability with chance distributions. Since the input parameters of the model contain fuzziness and randomness simultaneously, an algorithm based on chance measure is designed. The experimental results on the case application demonstrate the validity of the proposed method.

WOLAK A, ZAJĄC G, KUMBÁR V. **Evaluation of engine oil foaming tendency under urban driving conditions**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 229–235, http://dx.doi.org/10.17531/ein.2018.2.07.

The purpose of the article was to analyze the foaming tendency of engine oils used under excessive operating conditions. To achieve this end, foaming characteristics were determined for 23 oil samples in three measurement sequences. Foaming tendency was measured using the ASTM D 892 standard method, which consists in assessing foaming tendency of the liquid and foam stability. The cars used in the tests were uniform in terms of brand, type and operating conditions. The relationship between the mileage of the cars tested and the volume as well as stability of foam in used engine oils were presented using scatter plots with regression lines, correlation coefficient and 95% confidence interval. Based on the obtained results it was found that foaming tendency for new oils is characterized by high variability. The strongest foaming tendency at 24°C and 93°C (Sequence I and II) was observed for two out of five oil groups. Statistically significant differences were found between mileage and foaming tendency/foam stability for individual oils tested.

SONG H, SCHNIEDER E. **Modeling of railway system maintenance and availability by means of colored Petri nets**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 236–243, http://dx.doi.org/10.17531/ein.2018.2.08.

Prognostics and health management (PHM) technologies permit actionable information to enable proper decision-making for improving systems' performance. With the increasing requirements placed on the rail systems' availability, better maintenance decisions should be evaluated before practical application. The aim of this work is to build maintenance models and estimate the performance of considered maintenance

położenia środka ciężkości, jak również możliwej rekonstrukcji, którą przeprowadza się w celu lepszego dostosowania maszyny do warunków pracy.

ZHU L, ZHANG Y, ZHANG R, ZHANG P. **Zależna od czasu niezawodność układu przekładni zębatej jako funkcja procesu stopniowego zużycia**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 207–218, http://dx.doi.org/10.17531/ein.2018.2.05.

W artykule badano prawo dynamicznej ewolucji niezawodności mechanicznej powodowanej zużyciem na przykładzie układu przekładni zębatej. Na podstawie rozważań nad wpływem luzu powstałego na skutek zużycia na obciążenie zębów przekładni w obszarze podwójnych zazębień, wyprowadzono wzór na dynamiczny rozkład obciążeń przyjmowanych przez pary sąsiadujących zębów. Rozłożone naciski i prędkość zazębiania, które należy uwzględnić przy obliczaniu zużycia przekładni, otrzymano na podstawie modelu powierzchniowego Winklera oraz zasady zazębienia. W oparciu o model zużycia Archarda, wyprowadzono numeryczny model symulacyjny zużycia w przekładni zębatej oraz obliczono głębokość zużycia każdego z punktów zazębienia na zarysie zębów przy różnych cyklach zużycia. Aby uniknąć problemu czasochłonności, obliczeniowy model zużycia zastąpiono modelami zastępczymi bazującymi na sieci neuronowej i metodzie krigingu. Model procesu losowego zintegrowano z modelem zastępczym, a dynamiczną niezawodność dla nieliniowej struktury stochastycznej o nieznanej charakterystyce rozkładu uzyskano za pomocą techniki serii Edgeworth opartej na sieci neuronowej oraz metody czterech momentów, którą porównano z metodą symulacji Monte Carlo opartą na krigingu. Wykazano także wydajność obliczeniową i dokładność omawianej metody.

ZHANG L, ZHANG J, ZHAI H, ZHOU S. **Nowa metoda oceny niezawodności mechanizmów oparta na pomiarze szansy wystąpienia zdarzenia w warunkach niepewności rozmytej i losowej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 219–228, http://dx.doi.org/10.17531/ein.2018.2.06.

Tradycyjne metody analizy niezawodności oparte na teorii prawdopodobieństwa i teorii zbiorów rozmytych znajdują szerokie zastosowanie w praktyce inżynierskiej. Jednak metod tych nie można stosować do bezpośredniego pomiaru niepewności niezawodności przy niepewnych zmiennych, tj. subiektywnych zmiennych losowych i rozmytych. Aby zaradzić temu problemowi, przedstawiono nową metodę kwantyfikacji niezawodności opartą na teorii szansy, która jednocześnie spełnia aksjomaty dwoistości losowości oraz subaddytywności związanej z rozmytością w problemach niezawodności. Biorąc pod uwagę fakt, że systemy zazwyczaj charakteryzują się wielopoziomową strukturą, a uszkodzenia elementów składowych mają charakter wieloprzyczynowy, w niniejszym artykule zaproponowano model niezawodności eksploatacji systemu wielostanowego oparty na teorii szansy. W proponowanej metodzie, zamiast miar prawdopodobieństwa i możliwości, do kwantyfikacji niezawodności, w przypadku gdy dane są subiektywne zmienne losowe lub zmienne rozmyte, przyjęto miarę szansy wystąpienia zdarzenia. Do reprezentacji parametrów losowych i rozmytych wykorzystano zmienne hybrydowe, które stanowią podstawę dla wyprowadzenia rozwiązań w celu analizy niezawodności mechanizmu opartej na teorii szansy z rozkładem szans. Ponieważ parametry wejściowe modelu noszą jednocześnie znamiona rozmytości i losowości, opracowano algorytm oparty na mierze szansy. Wyniki eksperymentalne otrzymane na podstawie studium przypadku dowodzą poprawności proponowanej metody.

WOLAK A, ZAJĄC G, KUMBÁR V. **Ocena zmian charakterystyk pienienia olejów silnikowych użytkowanych w warunkach jazdy miejskiej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 229–235, http://dx.doi.org/10.17531/ein.2018.2.07.

Celem opracowania była analiza zmian charakterystyk pienienia olejów silnikowych użytkowanych w warunkach jazdy miejskiej. W związku z realizacją sformułowanego celu oznaczono skłonność do pienienia oraz trwałość piany dla 23 próbek oleju, według normy ASTM D 892. Próbki oleju pochodziły z samochodów stanowiących jednolitą flotę pod względem: marki, typu oraz warunków pracy silnika. Zależności przebiegu eksploatacyjnego badanych samochodów, z ilością i trwałością piany, w przepracowanych olejach silnikowych sprawdzono poprzez zastosowanie wykresów rozrzutu z linią regresji, współczynnikiem korelacji oraz 95%-owym przedziałem ufności. Na podstawie uzyskanych wyników stwierdzono, że skłonność do tworzenia piany dla olejów świeżych charakteryzuje się dużym zróżnicowaniem. Największą skłonność do pienienia w temp. 24°C i 93°C (Sekwencja I i II) zaobserwowano dla dwóch (z pięciu) grup olejowych. Potwierdzono istotne statystycznie różnice pomiędzy przebiegiem badanych pojazdów a poziomami poszczególnych charakterystyk pienienia.

SONG H, SCHNIEDER E. **Modelowanie utrzymania ruchu i gotowości systemu kolejowego za pomocą kolorowych sieci Petriego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 236–243, http://dx.doi.org/10.17531/ein.2018.2.08.

Technologie prognostyki i zarządzania zdrowiem (PHM) dostarczają praktycznych danych, które umożliwiają podejmowanie właściwych decyzji w zakresie poprawy wydajności systemów. Wraz z rosnącymi wymaganiami dotyczącymi gotowości systemów kolejowych, rośnie potrzeba oceny decyzji dotyczących utrzymania ruchu przed ich wprowadzeniem w życie. Celem przedstawionej pracy było zbudowanie modeli utrzymania ruchu oraz

nance decisions regarding the rail system's reliability and availability by means of Colored Petri nets. As a high-level formalization method, Colored Petri nets provide different color sets, which are suitable to represent different maintenance attributions. The maintenance models are evaluated at both the structure and parameterization levels. At the structure level, the structure correctness of the maintenance models is evaluated by using the state space analysis. At the parameterization level, specific maintenance decisions are illustrated. With various maintenance parameters, comparisons of system reliability and availability are made with the results obtained with the Colored Petri nets model.

YU H, ZHANG G, RAN Y, LI M, WANG Y. **A comprehensive and practical reliability allocation method considering failure effects and reliability costs**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 244–251, http://dx.doi.org/10.17531/ein.2018.2.09.

In view of the drawbacks in existing allocation methods which are incomplete considerations and poor practicality, a comprehensive fuzzy allocation method considering failure effects and reliability costs is proposed. Fuzzy linguistics and triangular fuzzy numbers are used to evaluate the uncertainty and subjective factors in allocation process. The traditional risk priority numbers (RPNs) are modified to overcome the shortages which are the same factor weights and equal difference of failure effects in original methods. State of the arts, components intricacy and working conditions are used to construct the reliability costs model, which solves the difficulties of costs statistics and avoids the sophisticated calculations which exist in current allocation methods. The relationship between reliability costs and potential risk of subsystem is studied and the value range of it is given in this paper. A case example is given to illustrative the scientificity and practicability of proposed allocation method.

FERDYNUS M, KOTEŁKO M, KRAL J. **Energy absorption capability numerical analysis of thin-walled prismatic tubes with corner dents under axial impact**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 252–259, http://dx.doi.org/10.17531/ein.2018.2.10.

The paper presents results of a parametric study into energy absorption capability of thin-walled square section columns with redrawn dents, subjected to axial impact compressive load. Thin-walled aluminum tubes with four dents in the corners were under investigation. The varying parameters were the dent's depth and distance of the dent to the base. The study was performed using Finite Element numerical code. Three crashworthiness indicators were examined: peak crushing force, crash load efficiency and stroke efficiency. The numerical results are shown in load-shortening diagrams, as well as diagrams and maps of crashworthiness indicators. It was found, that the main factor influencing a crushing mode and, subsequently, energy absorption capability, is a dent depth. The dent distance from the base is of less importance. Also a position of a dent, either at the bottom, or at the top base (the load application point) does not influence the crushing behavior significantly. For the deepest dents the relative increase of crash load efficiency (CLE) amounts 25% in comparison with the column without dents.

ZIEJA M, WAŻNY M, STĘPIEŃ S. **Outline of a method for estimating the durability of components or device assemblies while maintaining the required reliability level**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 260–266, http://dx.doi.org/10.17531/ein.2018.2.11.

The paper includes a probabilistic method for evaluating the durability of components and device assemblies which operate under the impact of destructive processes. As a result of these processes, wear that causes deterioration of their cooperation conditions occurs. It is assumed that a component operates reliably when the wear does not exceed the acceptable (limit) values. In mathematical terms, this method is based on a differential equation, after the transformation of which, it is possible to obtain the Fokker-Planck type partial differential equation. The specific solution of this equation allows for obtaining the density function of the probability wear in the normal distribution form. The paper presents two methods for determining the durability. The first one involves the application of the wear density function, and the second one consists in determining the probability density function of the time of reaching the acceptable state, and its use in order to determine the component or assembly durability. The paper presents a numerical example on the aircraft technology operation process.

ZARNIK MS, NOVAK F, PAPA G. **Sensors in proactive maintenance – a case of LTCC pressure sensors**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 267–272, http://dx.doi.org/10.17531/ein.2018.2.12.

Sensors are a vital component part of any process-controlled system. Even though designed to properly operate at required conditions within the whole lifetime, all sensors exhibit some level of drift with time. When selecting the sensors for implementation in a system proactive maintenance their ageing in specific operating

oszacowanie za pomocą kolorowych sieci Petriego możliwości realizacji rozważanych decyzji konserwacyjnych dotyczących niezawodności i gotowości systemu kolejowego. Kolorowe sieci Petriego to metoda o wysokim poziomie formalizacji, którą w przedstawionej pracy wykorzystano do reprezentacji za pomocą różnych zestawów kolorów, różnych atrybutów utrzymania ruchu. Modele utrzymania ruchu oceniano zarówno na poziomie struktury jak i parametryzacji. Na poziomie struktury, poprawność struktury modeli utrzymania ruchu oceniano za pomocą analizy przestrzeni stanów. Na poziomie parametryzacji, zilustrowano konkretne decyzje dotyczące konserwacji. Niezawodność i gotowość systemu przy różnych parametrach utrzymania ruchu porównano z wynikami uzyskanymi za pomocą modelu kolorowych sieci Petriego.

YU H, ZHANG G, RAN Y, LI M, WANG Y. **Kompleksowa i praktyczna metoda alokacji niezawodności uwzględniająca skutki uszkodzeń i koszty niezawodności**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 244–251, http://dx.doi.org/10.17531/ein.2018.2.09.

Ze względu na niedostatki istniejących metod alokacji, które nie dają pełnego obrazu problematyki i mają słabe zastosowanie w praktyce, w artykule zaproponowano kompleksową metodę alokacji opartą na logice rozmytej, uwzględniającą skutki uszkodzeń i koszty niezawodności. W pracy wykorzystano lingwistykę rozmytą i trójkątne liczby rozmyte do oceny niepewności i czynników subiektywnych w procesie alokacji. Zmodyfikowano tradycyjny wskaźnik liczby priorytetowej ryzyka (RPN), co pozwoliło na poprawę mankamentów charakteryzujących oryginalną metodę, t.j. takie same współczynniki wagowe i równoważność skutków uszkodzeń o różnym stopniu ciężkości. Na podstawie wiedzy o stanie techniki, złożoności komponentów i warunkach pracy, skonstruowano model kosztów niezawodności, który rozwiązuje trudności dotyczące sporządzania statystyki kosztów i pozwala uniknąć skomplikowanych obliczeń stosowanych w obecnych metodach alokacji. Zbadano związek między kosztami niezawodności a potencjalnym ryzykiem podsystemu, oraz podano jego zakres wartości. Prezentowane studium przypadku demonstruje możliwe zastosowania i efektywność proponowanej metody.

FERDYNUS M, KOTEŁKO M, KRAL J. **Numeryczna analiza energochłonności cienkościennych słupów pryzmatycznych z przetłoczeniami**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 252–259, http://dx.doi.org/10.17531/ein.2018.2.10.

W artykule przedstawiono wyniki badań numerycznych zdolności pochłaniania energii energoabsorberów w postaci cienkościennych słupów o przekroju kwadratowym z wgłębieniami, poddanych osiowym obciążeniom udarowym. Badano wpływ parametrów geometrycznych oraz położenia inicjatorów zgniotu w postaci walcowych przetłoczeń w narożach na zachowanie się konstrukcji oraz właściwości energoabsorbcyjne (współczynnik efektywności zgniotu- Ste oraz procentowy stosunek siły średniej do maksymalnej - CLE). Obliczenia numeryczne prowadzono z wykorzystaniem MES, programu Abaqus 6.14. Wyniki przedstawiono w postaci charakterystyk obciążenie – skrócenie oraz diagramów i wykresów. Stwierdzono, że istotny wpływ na zachowanie się konstrukcji podczas uderzenia oraz jej energochłonność ma głębokość przetłoczenia, mniej istotne jest jego położenie. W przypadku słupów z najgłębszymi przetłoczeniami względny wzrost współczynnika CLE, w porównaniu z wynikami uzyskanymi dla słupa gładkiego wynosi 25%.

ZIEJA M, WAŻNY M, STĘPIEŃ S. **Zarys metody szacowania trwałości elementów lub zespołów urządzeń z zachowaniem wymaganego poziomu niezawodności**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 260–266, http://dx.doi.org/10.17531/ein.2018.2.11.

Praca zawiera probabilistyczną metodę oceny trwałości elementów lub zespołów urządzeń pracujących w warunkach oddziaływania procesów destrukcyjnych. W wyniku działania tychże procesów następuje zużywanie powodujące pogorszenie warunków ich współpracy. Przyjmuje się, że element pracuje niezawodnie, gdy zużycie nie przekracza wartości dopuszczalnych (granicznych). Metoda od strony matematycznej bazuje na równaniu różniczowym z którego po przekształceniu otrzymuje się równanie różniczkowe cząstkowe typu Fokkera-Plancka. Z rozwiązania szczególnego tego równania otrzymuje się funkcję gęstości prawdopodobieństwa zużywania w postaci rozkładu normalnego. W pracy przedstawione są dwa sposoby wyznaczania trwałości. Pierwszy polega na wykorzystaniu funkcji gęstości zużywania a drugi na wyznaczeniu funkcji gęstości prawdopodobieństwa czasu osiągania stanu dopuszczalnego i zastosowanie jej do wyznaczenia trwałości elementu lub zespołu. W pracy przedstawiono przykład liczbowy dotyczący procesu eksploatacji techniki lotniczej.

ZARNIK MS, NOVAK F, PAPA G.**Czujniki stosowane w konserwacji proaktywnej – przypadek ceramicznych czujników ciśnienia wykonanych w technologii LTCC**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 267–272, http://dx.doi.org/10.17531/ein.2018.2.12.

Czujniki stanowią istotny komponent każdego systemu kontrolowanego przez proces. Chociaż czujniki są zaprojektowane tak, aby prawidłowo działały w wymaganych warunkach w całym okresie eksploatacji, wszystkie wykazują jednak pewien poziom dryfu w czasie. Wybierając czujniki do wdrożenia w proaktywnej konserwacji systemu, należy

conditions should be considered as an important issue. Here we focus on thick-film piezoresistive sensors based on low temperature cofired ceramic (LTCC) and discuss their ageing in different regimes of operations. Frequent overloading and particularly with limit overpressures can result in observable drifts and unacceptable scattering from the calibrated characteristics. For the sensors operating in the water the overloads are even more critical. Moreover, under the regime with frequent overloads, some non-critical, intrinsic defects in the sensing structure, which normally do not affect the characteristics and are non-detectable by the output tests in serial production may develop into critical defects that shorten the sensor lifetime.

PENG R. **Reliability of interdependent networks with cascading failures**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 273–277, http://dx.doi.org/10.17531/ein.2018.2.13.

The reliability of network systems of various structures has been studied by many researchers. However, most of the works just consider the reliability of a single network system. In practice, different networks may be interdependent such that the failure in one network may result in the failure in another network. The cascading failures have been shown to be catastrophic by some researchers. However, the quantitative evaluation for the reliability of interdependent networks has not been proposed. In this paper, a multi-valued decision diagram based approach is presented to evaluate the reliability of interdependent networks. Illustrative examples are proposed to demonstrate the application of the framework.

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.

Apart from reliability evaluation, the methodology of network systems reliability assessment presented in the article enables the design of modernisation of such systems targeted mainly at ensuring their required reliability. In practice the methodology can be applied for various network systems, e.g. computer, power, gas, water distribution, telecommunications and transport networks. A reliability analysis of a transport network in urban public transport is presented. Calculations were performed for selected criteria of network availability which actually conditions the quality of transport services provided. The basic calculation tool used was the factoring algorithm that enabled the assessment of the impact of individual connections failure (in particular those caused by physical factors) on the reliability of the whole network. The feasibility of modernisation of the network analysed is discussed and the results are presented in diagrams.

PIEGDOŃ I, TCHÓRZEWSKA-CIEŚLAK B, EID M. **Managing the risk of failure of the water supply network using the mass service system**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 284–291, http://dx.doi.org/10.17531/ein.2018.2.15.

The aim of this paper is to analyse the functioning of the repair brigades in the process of failure removal in the water distribution subsystem using the mass service system (MSS). An example is presented using queuing model which takes into account notifications with various scheduling algorithms of failures occurring to the system. The functioning analysis of mass service system can be used in the optimization of the repair teams' actions and in the management of water supply companies.

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.

Operational vehicle quality measures are an important element used to evaluate the performance of transport services. In practice, there are many methods involved in the operational evaluation of vehicles. They are characterized in this article. Artificial Intelligence methods, especially artificial neural networks, can also be successfully used for this purpose, and especially when deciding on quality assessment processes for machines, including motor vehicles. The use of methods to support decision-making based on facts is extremely important for the credibility and objectivity of the evaluation. These methods can also be used in relation to the use of vehicles in the assessment of transport services. The article presents the method of using artificial neural networks for the operational evaluation of vehicles used in freight transport services. The basis for the verification of the method was an experimental research carried out at a company making dairy products, cooperating with transport companies, supplying products for the production process. The results obtained from the operation of vehicles from the studied companies have confirmed, at the probability level of 99%, high efficiency of the proposed method in evaluating transport services using operational vehicle quality measures.

koniecznie rozważyć ich starzenie się w określonych warunkach pracy. Przedmiotem artykułu są grubowarstwowe czujniki piezorezystancyjne wykonane z ceramiki technologią LTCC (niskotemperaturowej ceramiki współwypalanej) oraz ich starzenie się w różnych trybach działania. Częste przeciążanie, zwłaszcza przy nadciśnieniu granicznym, może powodować dryf i niedopuszczalny rozrzut wskazań w stosunku do charakterystyk wzorcowych. W przypadku czujników pracujących w wodzie, przeciążenia mają jeszcze bardziej krytyczny charakter. Co więcej, w trybie pracy z częstymi przeciążeniami, niektóre niekrytyczne wady wewnętrzne w strukturze sensorowej, które normalnie nie mają wpływu na charakterystykę czujnika i są niewykrywalne w badaniach kontrolnych wyrobu gotowego w produkcji seryjnej, mogą przeobrażać się w wady krytyczne, które skracają cykl życia czujnika.

PENG R. **Niezawodność współzależnych sieci z uszkodzeniami kaskadowymi**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 273–277, http://dx.doi.org/10.17531/ein.2018.2.13.

Niezawodność systemów sieciowych o różnych strukturach stanowi przedmiot licznych badań. Jednak większość prac dotyczy tylko niezawodności pojedynczych systemów sieciowych. W praktyce, różne sieci mogą działać współzależnie, tak iż awaria jednej może powodować awarię innej sieci. Niektóre badania pokazują, że uszkodzenia kaskadowe są uszkodzeniami katastroficznymi. Nie zaproponowano jednak dotąd ilościowej oceny niezawodności współzależnych sieci. W niniejszym artykule przedstawiono podejście oparte na koncepcji wielowartościowego diagramu decyzyjnego, które pozwala na ocenę niezawodności wzajemnie zależnych sieci. Przedstawiono przykłady ilustrujące zastosowanie proponowanego paradygmatu.

MŁYNARSKI S, PILCH R, SMOLNIK M, SZYBKA J. **Metodyka szacowania niezawodności układów sieciowych na przykładzie komunikacji miejskiej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 278–283, http://dx.doi.org/10.17531/ein.2018.2.14.

W artykule zaprezentowano opracowaną metodykę szacowania niezawodności układów sieciowych. Rozwiązanie to umożliwia dokonywanie oceny niezawodności oraz projektowanie modernizacji rozpatrywanej sieci przede wszystkim w aspekcie zapewnienia jej wymaganej niezawodności. Praktyczne wykorzystanie omawianej metodyki może mieć miejsce w odniesieniu do różnych układów sieciowych, np. sieci komputerowych, energetycznych, gazowych, wodociągowych, telekomunikacyjnych i transportowych. W artykule przedstawiono analizę niezawodności sieci komunikacyjnej w miejskim transporcie zbiorowym. Obliczenia przeprowadzono dla wybranych kryteriów zdatności sieci, które praktycznie warunkują jakość świadczonych usług transportowych. Podstawowe narzędzie obliczeniowe stanowił algorytm faktoryzacji, który umożliwia ocenę wpływu uszkodzeń poszczególnych połączeń (spowodowanych w szczególności czynnikami fizycznymi) na niezawodność całej sieci. W opracowaniu uwzględniono możliwość modernizacji analizowanej sieci, a uzyskane wyniki przedstawiono na wykresach.

PIEGDOŃ I, TCHÓRZEWSKA-CIEŚLAK B, EID M. **Zarządzanie ryzykiem awarii sieci wodociągowej z wykorzystaniem system masowej obsługi**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 284–291, http://dx.doi.org/10.17531/ein.2018.2.15.

Celem pracy jest analiza funkcjonowania brygad naprawczych w procesie usuwania awarii w podsystemie dystrybucji wody przy użyciu systemu masowej obsługi (SMO). Przykład został przedstawiony przy użyciu modelu kolejek, który uwzględnia zgłoszenia napływające do systemu z różnymi algorytmami planowania awarii. Funkcjonująca analiza systemu masowego świadczenia usług może być wykorzystana w optymalizacji działań zespołów naprawczych oraz w zarządzaniu przedsiębiorstwem wodociągowym.

ŚWIDERSKI A, JÓŹWIAK A, JACHIMOWSKI R. **Eksploatacyjne miary jakości pojazdów w zastosowaniu do oceny usług transportowych z wykorzystaniem sztucznych sieci neuronowych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 292–299, http://dx.doi.org/10.17531/ein.2018.2.16.

Eksploatacyjne miary jakości pojazdów są istotnym elementem wykorzystywanym do oceny realizacji usług transportowych. W praktyce mamy do czynienia z wieloma metodami związanymi z eksploatacyjną oceną pojazdów. Scharakteryzowano je w artykule. Metody sztucznej inteligencji, a zwłaszcza sztuczne sieci neuronowe, również mogą być z powodzeniem wykorzystane do tego celu, a zwłaszcza przy podejmowaniu decyzji w procesach oceny jakości maszyn, w tym pojazdów samochodowych. Zastosowanie metod, które pozwalają wspomagać proces decyzyjny na podstawie faktów jest niezmiernie istotne z punktu widzenia wiarygodności i obiektywności oceny. Metody te mogą być również wykorzystane w odniesieniu do eksploatacji pojazdów w zastosowaniu do oceny usług transportowych. W artykule przedstawiono metodę wykorzystania sztucznych sieci neuronowych do eksploatacyjnej oceny pojazdów wykorzystywanych w usługach transportowych towarów. Podstawę weryfikacji metody stanowiły badania eksperymentalne przeprowadzone w przedsiębiorstwie produkującym produkty mleczarskie, współpracującym z firmami transportowymi, dostarczającymi wyroby do produkcji. Uzyskane wyniki potwierdziły z 99-procentowym prawdopodobieństwem wysoką skuteczność proponowanej metody w dokonywaniu oceny usług transportowych z wykorzystaniem eksploatacyjnych miar jakości pojazdów.

DONG F, LIU Z, WU Y, HAO J. **A multi-stage risk-adjusted control chart for monitoring and early-warning of products sold with two-dimensional warranty**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 300–307, http://dx.doi.org/10.17531/ein.2018.2.17.

Warranty claims data contain valuable information about the quality and reliability of products. The monitoring and early-warning of warranty claims data are of great significance to the manufacturer by identifying and solving the emerging quality or reliability problem as soon as possible. However, though it has been used widely in the automobile industry, there are no studies that have been carried out on the monitoring and early-warning of claims data for products sold with two-dimensional warranty. In order to fill this gap, fitting the two-dimensional warranty claims data with accelerated failure model (AFT), a multi-stage risk-adjusted control chart is proposed by this paper, for which a reasonable product sales tracking time and a monitoring time are suggested to reduce the influence of sales delay and fluctuating claim rates. Comparing with traditional Cumulative Sum Control Chart (CUSUM), the applicability and availability of the proposed model are demonstrated in the final.

DONG F, LIU Z, WU Y, HAO J. **Karta kontrolna do wieloetapowego monitorowania produktów sprzedawanych z gwarancją dwuwymiarową, z korektą ryzyka i wczesne ostrzeganie o wadach produkcyjnych na podstawie danych z reklamacji**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 300–307, http://dx.doi.org/10.17531/ein.2018.2.17.

Roszczenia gwarancyjne stanowią cenne źródło informacji na temat jakości i niezawodności produktów. Monitorowanie danych dotyczących roszczeń gwarancyjnych i wczesne ostrzeganie w oparciu o te dane ma wielkie znaczenie dla producenta, ponieważ pozwala rozpoznawać i rozwiązywać pojawiające się problemy związane z niezawodnością w jak najkrótszym czasie. Chociaż ten rodzaj monitorowania i wczesnego ostrzegania jest szeroko stosowany w przemyśle motoryzacyjnym, nie przeprowadzono dotąd żadnych badań na temat tych procesów w odniesieniu do produktów sprzedawanych z gwarancją dwuwymiarową. W celu wypełnienia tej luki, dane o reklamacjach składanych na podstawie gwarancji dwuwymiarowych dopasowano modelem uszkodzeń przyspieszonych (accelerated failure model, AFT), a następnie przedstawiono koncepcję karty kontrolnej monitorowania wieloetapowego z korektą ryzyka, dla której zaproponowano odpowiedni czas śledzenia sprzedaży produktu i czas monitorowania, mając na uwadze zmniejszenie wpływu opóźnień w sprzedaży i wahań liczby roszczeń zgłaszanych z tytułu gwarancji. Możliwości zastosowania i dostępność proponowanego modelu porównano z tradycyjną kartą sum skumulowanych.

KILIAN-BŁAŻEJEWSKA E. **The influence of selected factors on the failure rate of the water pipelines located on the area of the impact of mining tremors**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 308–317, http://dx.doi.org/10.17531/ein.2018.2.18.

The article presents the influence of the selected factors, including mining tremors, described by the parameter PGVHmax on the failure rate of the water distribution pipelines. In created multiple regression models, the following independent variables were used: diameter and material from which the pipes were made, operation time without failure, the age of pipes, the value of pressure and PGVHmax. The values of PGVHmax in places with damaged water pipelines were determined by constructed the seismic wave propagation models. The analysis was carried out for a random sample of all water pipelines and extracted from this sample new groups: steel and gray cast iron pipes, their diameters, diameters and materials, and their construction time.

KILIAN-BŁAŻEJEWSKA E. **Wpływ wybranych czynników na intensywność uszkodzeń przewodów podsystemu dystrybucji wody, zlokalizowanych na terenie oddziaływania wstrząsów pochodzenia górniczego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 308–317, http://dx.doi.org/10.17531/ein.2018.2.18.

W artykule przedstawiono badania wpływu wybranych czynników, w tym wstrząsów pochodzenia górniczego, opisanych za pomocą parametru PGVHmax na intensywność uszkodzeń przewodów podsystemu dystrybucji wody. Jako zmienne niezależne w utworzonych modelach regresji wielorakiej przyjęto: średnicę i materiał z którego wykonane są przewody, czas pracy bezuszkodzeniowej, wiek przewodów, wysokość ciśnienia i PGVHmax. Wartości PGVHmax w miejscach występowania awarii przewodów wodociągowych, wyznaczone zostały na podstawie zbudowanych modeli propagacji fali drgań w ośrodku gruntowym. Analiza przeprowadzona została dla próby losowej obejmującej sumarycznie wszystkie przewody sieci wodociągowej magistralnej, rozdzielczej i przyłącza oraz dla wyodrębnionych z tej grupy prób losowych obejmujących: przewody zbudowane ze stali i z żeliwa szarego, przewody z uwzględnieniem ich średnicy, średnicy i materiału oraz z uwzględnieniem czasu ich budowy.

WEREMCZUK A, BOROWIEC M, RUDZIK M, RUSINEK R. **Stable and unstable milling process for nickel superalloy as observed by recurrence plots and multiscale entropy**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 318–326, http://dx.doi.org/10.17531/ein.2018.2.19.

This paper discusses the stability of high-speed machining processes. The problem of harmful vibrations can usually be detected based on measured signal forces. Nevertheless, the chatter effect may be unrevealed and hence some alternative approaches of signal monitoring must be taken to detect it. In the discussed case of machining, process stability is determined by means of stability diagrams. The measured milling force components are investigated by various signal analysis methods. In addition to this, the analysis also uses recurrence plots, recurrence quantifications, composite multi-scale-entropy and as well the statistical approach. Results obtained by the different methods are presented and discussed.

WEREMCZUK A, BOROWIEC M, RUDZIK M, RUSINEK R. **Stabilny i niestabilny proces frezowania super stopu niklu obserwowany z wykorzystaniem wykresów rekurencyjnych i entropii wieloskalowej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 318–326, http://dx.doi.org/10.17531/ein.2018.2.19.

W niniejszym artykule omówiono stabilność procesu obróbki szybkościowej. Problem szkodliwych drgań zwykle może zostać wykryty na podstawie sygnału mierzonych sił. Niemniej jednak drgania samowzbudne mogą nie ujawnić się w sposób wyraźny, a niekiedy do ich wykrycia potrzebne jest alternatywne podejście do monitorowania sygnału. W przedstawionym procesie obróbki stabilność procesu oszacowano za pomocą wykresów stabilności. Zmierzone siły frezowania badano różnymi metodami analizy sygnału. W analizie wykorzystano wykresy rekurencyjne, wskaźniki rekurencyjne, entropię wieloskalową, a także podejście statystyczne. Przedstawiono wyniki różnych metod i omówiono ich porównanie.

BABISHIN V, HAJIPOUR Y, TAGHIPOUR S. **Optimisation of non-periodic inspection and maintenance for multicomponent systems**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 327–342, http://dx.doi.org/10.17531/ein.2018.2.20.

A *k*-out-of-*n*:G system and a system with components subject to soft and hard failures are both inspected non-periodically. For the *k*-out-of-*n* system, components fail "silently" (i.e. are hidden), and the entire system fails when (*n*-*k*+1)st component fails. For the system with hard-type and soft-type components, hard failures cause system failure, while soft failures are hidden and do not cause immediate failure of the system, but still reduce its reliability. Every system failure allows for an opportunistic inspection of hidden soft-type components in addition to the scheduled inspections. The available maintenance types are replacement and minimal repair. For hard-type components, the maintenance decision is determined by the optimal age before replacement. For the soft-type components with hidden failures, we do not know their age, and so decide on the appropriate type of maintenance using the optimal number of minimal repairs before replacement. The hidden nature of soft-type component failures precludes the use of a tractable analytic expression, so we use simulation and genetic algorithm (GA) to jointly optimise the non-periodic policies on maintenance and inspection and to ensure these incur minimal expected total cost over a finite planning horizon. Due to increasing computational complexity associated with the number of inspections and maintenance policies to be evaluated,

BABISHIN V, HAJIPOUR Y, TAGHIPOUR S. **Optymalizacja nie-okresowych przeglądów i konserwacji systemów wieloelementowych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 327–342, http://dx.doi.org/10.17531/ein.2018.2.20.

Przeglądów systemu typu *k* z *n*: G oraz systemu z elementami ulegającymi miękkim i twardym uszkodzeniom dokonuje się nie-okresowo. W przypadku systemu *k* z *n*, składowe ulegają uszkodzeniom „w trybie cichym" (tj. uszkodzenia są ukryte), a cały system ulega awarii, gdy ulegnie uszkodzeniu element (*n*-*k* + 1). W przypadku systemu z elementami typu twardego i miękkiego, uszkodzenia twarde prowadzą do awarii systemu, natomiast uszkodzenia miękkie są ukryte i nie powodują natychmiastowej awarii systemu, choć nadal zmniejszają jego niezawodność. Każda awaria systemu stanowi dodatkową, w stosunku do przeglądów planowych, okazję do przeprowadzenia przeglądu (tzw. przegląd awaryjny) ukrytych elementów miękkich. Dostępne rodzaje konserwacji to wymiana oraz naprawa minimalna. W przypadku komponentów twardych, decyzję, który typ konserwacji zastosować, podejmuje się biorąc pod uwagę optymalny wiek przed wymianą. W przypadku elementów miękkich z ukrytymi uszkodzeniami, wiek optymalny jest nieznany, dlatego decyzje o odpowiednim typie konserwacji podejmuje się z uwzględnieniem optymalnej liczby minimalnych napraw przed wymianą. Ukryty charakter uszkodzeń elementów składowych typu miękkiego wyklucza wykorzystanie rozwiązywalnego wyrażenia analitycznego, dlatego w pracy użyto symulacji i algorytmu genetycznego (GA), w celu jednoczesnej optymalizacji nieokresowych strategii prowadzenia konserwacji i przeglądów oraz zapewnienia, że

the genetic algorithm presents a promising method of optimisation for complex multicomponent systems with multiple decision parameters.

będą one pociągały za sobą minimalny oczekiwany koszt całkowity w skończonym horyzoncie planowania. W świetle rosnącej złożoności obliczeniowej związanej z dużą liczbą ocenianych przeglądów i strategii utrzymania ruchu, algorytm genetyczny stanowi obiecującą metodę optymalizacji złożonych systemów wieloelementowych o wielu parametrach decyzyjnych.

Article citation info:
HAWRYLUK M, GRONOSTAJSKI Z, ZIEMBA J, DWORZAK Ł, JABŁOŃSKI P, RYCHLIK M. Analysis of the influence of lubrication conditions on tool wear used in hot die forging processes. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 169–176, http://dx.doi.org/10.17531/ein.2018.2.01.

Marek HAWRYLUK
Zbigniew GRONOSTAJSKI
Jacek ZIEMBA
Łukasz DWORZAK
Paweł JABŁOŃSKI
Marcin RYCHLIK

ANALYSIS OF THE INFLUENCE OF LUBRICATION CONDITIONS ON TOOL WEAR USED IN HOT DIE FORGING PROCESSES

ANALIZA WPŁYWU WARUNKÓW SMAROWANIA NA ZUŻYCIE NARZĘDZI W PROCESACH KUCIA MATRYCOWEGO NA GORĄCO

The study discusses the subject of lubrication in the processes of hot die forging with the consideration of the durability of forging tools and instrumentation. It presents a literature research as well as the authors' own investigations of the effect of the use of cooling and lubrication agents, the amount of the dosage as well as the direction of its application, and also the factors influencing the tribological conditions. The lubrication devices and systems currently applied in the industry have been analyzed as well. On this basis, making use of their knowledge and experience, the authors have developed and constructed a lubricating device. The elaborated system, implemented into the industrial process, makes it possible to select and ensure the optimal tribological conditions of the process by way of controlling the amount and frequency of the applied lubricant dosage. It can constitute an alternative for the manual method of lubricant application, which is dependent on the human factor, or the fully automated, yet expensive, lubrication systems. The obtained test results point to potential possibilities of a permanent introduction of the constructed device also into other forging processes, through its integration with a manipulator. The proposed solution ensures more stability and higher repeatability of the lubrication conditions as well as increases the efficiency of the production process, thus significantly reducing the unit costs of the production of forgings.

Keywords: *hot die forging, lubricating system, wear, tribological conditions.*

Praca dotyczy problematyki smarowania w procesach kucia matrycowego na gorąco z uwzględnieniem trwałości narzędzi i oprzyrządowania kuźniczego. Przedstawiono badania literaturowe oraz własne autorów dotyczące wpływu zastosowania środków smarno-chłodzących, ilości dawki i kierunku jej podawania oraz innych czynników wpływających na warunki tribologiczne. Przeanalizowano także obecnie stosowane w przemyśle urządzenie i systemy smarowania. Na tej podstawie autorzy w oparciu o wiedzę i doświadczenie opracowali i zbudowali urządzenie smarujące. Opracowany system, zaimplementowany do przemysłowego procesu pozwala na dobór i zapewnienie optymalnych warunków tribologicznych w procesie poprzez sterowanie ilością i częstotliwością podawanej dawki środka smarnego. Może być ono alternatywą dla manualnej metody nanoszenia środka smarnego, zależnej od czynnika ludzkiego lub w pełni zautomatyzowanych, lecz drogich systemów smarowania. Uzyskane wyniki badań wskazują na potencjalne możliwości wprowadzenia na stałe do pracy zbudowanego urządzenia także do innych procesów kucia poprzez integrację z manipulatorem. Zaproponowane rozwiązanie zapewnia większą stabilność i powtarzalność warunków smarowania oraz pozytywnie wpływa na zwiększenie wydajności procesu wytwarzania, a tym samym znacząco obniża jednostkowe koszty produkcji odkuwek.

Słowa kluczowe: *kucie matrycowe na gorąco, system smarujący, zużycie, warunki tribologiczne.*

1. Introduction

During the processes of die forging performed at elevated and high temperatures, the tools are subjected to very large, periodical thermal loads from 80 oC to 600 oC, as well as to mechanical loads reaching above 1200 MPa [39]. The primary and most common destructive mechanisms are [13, 27]: abrasive wear [4], mechanical cracking, plastic strain [23] as well as thermal [9, 36] and thermo-mechanical fatigue [7]. The progression of the loading exerted on the dies and punches is periodical in nature. It is a combination of thermal and mechanical loads arising from the contact with and deformation of the hot preform by the cold tool. A variable thermal load has a particular impact on the die's lifetime. It is the leading cause of fatigue crack formation and changes in the physical and mechanical proper-

ties of the tool's surface layer [3, 32]. At the same time, thermal load intensifies the abrasive wear caused by high mechanical loads, and this translates to a shorter operating life of forging tools and accessories [25, 26].

Lubrication is critical to the lifetime of forging tools. In the standard operating procedure during a semi-closed die forging process (upsetting, flattening, etc.), tool lubrication is not utilized as the flow of the material being deformed is uncomplicated. This is not the case in typical die forging, during which the use of a lubricating and cooling agent is required due to the need to minimize the friction for the purpose of precise filling of the tool's impression with the material. Above all, the application of lubricating and cooling agents reduces the friction between the forged material and the tool material, and insulates the tool's material (impression) against a direct contact with the hot forged material. This effectively reduces the die's surface temperature, thus reducing the intensity of the tempering, oxidation and erosion processes [13]. A flaw of the use of lubricating and cooling agents is the sudden cooling of the surface layer, which may accelerate the thermal fatigue process. In addition, the lubricant also reduces the friction coefficient after the forged metal is released from the die [1, 2, 4]. Not only the properties of the lubricant, but also the method and direction of the feed, size and frequency of the lubricant dosage, play a role in proper lubrication. It is also important to ensure a repeatable and uniform distribution of the lubricant. Currently, non-automated, and providing low repeatability, manually operated lubricating devices are still commonly used in many die forges. For this reason, the application of more or less automated lubricating and cooling devices and systems is currently being adopted more and more frequently, as they enable precise lubricant dosing [12, 18, 22]. The proper lubrication does not only determine whether a forging without defects, such as underfills, will be produced, but also has the benefit of reducing the tool wear [20]. As automation develops, automated lubricating devices and systems are created [5, 11, 17]. Such solutions are easy to use, and have been implemented successfully in German forges, leaders in the automation of production. In addition, owing to the application of manipulators, it has become possible to build a fully flexible system, making it possible to control all the significant lubrication parameters, such as the nozzle position, the time of application, the proportions of substances, etc. [10, 18]. Additionally, such systems are synchronized with the work of the forging unit, which, according to this method of lubricant feeding, eliminates the subjective factor of an experienced human operator [19]. However, this is linked to rather high costs of investment, which may not be balanced out if only the profit resulting from raising the product quality and the tool lifetime are considered.

This is why an intensive search is currently underway for new solutions concerning both lubricants dedicated for specific forging applications and modern automated lubricating devices ensuring optimal friction conditions, which, at elevated temperatures, are even more critical to hot and semi-hot industrial die forging processes. Further research and development work in this field is completely justified from both the scientific and financial perspective, as the problem of effective lubrication continues to be unsolved and poses a significant challenge to many research centers and industrial enterprises.

The aim of the paper is to analyze the results of studies concerning the effect of the amount, frequency and application manner of the lubricating agent, as well as the use of a proprietary cooling and lubrication system, on the durability of the forging instrumentation in hot die forging processes.

2. Analysis of the state of the problem and research of factors influencing the operation of tools in die-forging processes

It should be emphasized that there is a lot of space devoted to the research and analysis of the properties of lubricants used in hot forging processes in the available literature [1, 11, 12, 21, 24, 35]. In addition, a continuous technological development has led to an increase, outside the research centers, in studies and industrial applications involving lubricants dedicated to specific processes on the very competitive lubricants market. The leading manufacturers of lubricants in the world include such companies as:Acheson, Fuchs, Henkel/Bechem, Houghton, Oelheld, etc., as well as Naftochem from Poland.

In addition to the undisputed role of lubricants, the important factors influencing the lubrication are the appropriate method of applying the lubricant technology as well as the lubricant dosage. Currently, the most popular method of lubrication is the use of a direct, manual spray lubricant performed by the blacksmith during the process. Its main flaw is a non-uniform distribution of the lubricant, which causes a non-unifrom die temperature distribution and, in effect, a faster local wear of the tool. Only in the case of transfer presses, fully automated systems and lubricating devices are used, synchronized with the work of the whole aggregate.

In the vast majority of works in the subject literature, the most space is devoted to lubrication systems in cold forging processes [31], mainly for aluminum [33], or cold and warm processes [5, 6, 28], while there is definitely less focus on the case of systems and devices destined for warm and hot forging [29, 34, 37]. Even in the extensive review paper by Altan's [1], you can only find brief information about lubrication systems, which is more related to the selection of lubricants than lubricating devices. On the other hand, in the work [30], you can find information about the devices and systems equipped with the forging sets used in the Japanese forging industry, but also in the quoted work, there is no data on lubrication systems. Also interesting are the results of the European Research Project "Brite-Euram" [31] on the development of environmentally friendly systems for the lubrication of tools in the hot-forging of steel. For example, in the paper [2], in a study conducted by a German-Japanese team based on experiments during production, guidelines were developed, such as a

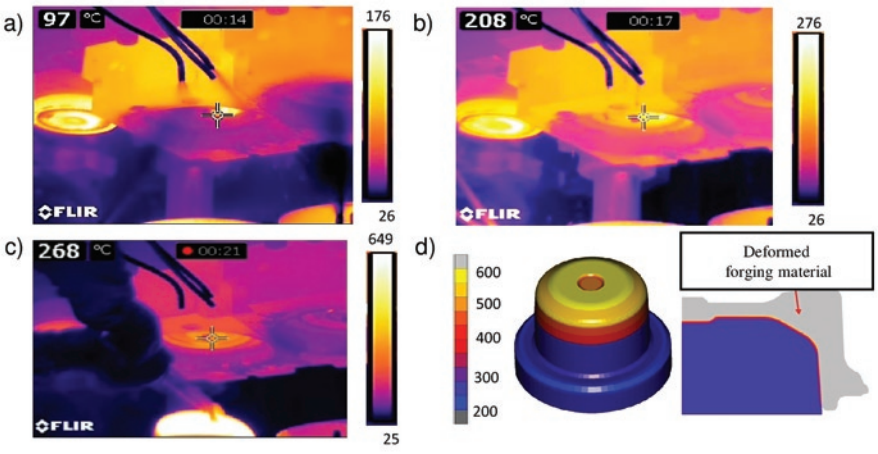


Fig. 1. Thermograms with temperature distributions on top punches used in the process of forging a disk:a) tool lubrication cycle – the temperature at the marked point is 97°C, b) 3 seconds after the end of lubrication – temperature increase by over 100°C, c) just after the forging process – about 1 second after deformation, d) temperature field on the analyzed punch at the time of forging (contact) determined by numerical modeling

scheme for the selection of the optimal lubrication systems to increase durability of the forging tools. In turn, the study [20] presents selected factors affecting the quality of lubrication and tool wear. The work [35] is also noteworthy, in which the authors presented the possibilities of modeling the direction and amount of lubricant dosage.

As shown in the introduction, the processes of hot forging is performed under extreme operating conditions, which means that they are among the most difficult manufacturing processes to carry out. The authors of the article have been conducting numerous studies of the impact of the tribological conditions on the exploitation of forging tools for many years now [15, 16]. For example, Fig. 1 presents the results of the research on the influence of the lubricant application on the temperature changes on the tool surface during the forging cycle.

As can be observed on the basis of the thermograms taken by a thermal imager (Fig. 1a-c), the temperature on the tool's surface at a selected point is approx. 100°C (Fig.1a) over the course of only about 7 seconds of the forging process, at the time of the final lubrication phase, which lasts about 2-3s (Fig. 1a). After the next 3 seconds, the temperature rises rapidly up to the mean "working" temperature of the tool in the forging process (Fig. 1b). Just after the forging – deformation – over the next 4 seconds – the temperature of the tool increased by 60°C (Fig. 1c) as a result of the conversion of the strain work to heat.

Research employing numerical modeling was also conducted, showing that, at the time of the contact between the formed material and the punch, the temperature increases to over 500°C (Fig. 1d). Such dynamic temperature changes on the tools' working surfaces may significantly reduce their lifetime in the absence of properly selected lubrication parameters, due to excessive periodical overcooling or non-targeted feeding of the lubricant dose, which is unsynchronized with the work of the forging unit. It seems that even the most aptly selected lubricant will not have a significant impact on the temperature distributions.

Fig. 2a presents a standard manner of lubrication used in crank presses, which is a manual adjustment of the lubricating nozzles, according to the decision and experience of the operator.

The lack of sufficient control over the lubrication process and the inappropriate adjustment of the nozzles at the beginning of the process are the causes of a non-uniform wear (Fig.2b and Fig. 2c), which, in turn, causes a more intensified damage, shortening the tool operation time, and in the case of a long operation, possibly affecting the shape of the forging.

Do not synchronize the exposure of the lubricant to the work of the forging press and/or forging transfer for subsequent tools (Fig.3a). An inappropriately selected dose of the cooling and lubricating agent can be the cause of premature wear of the tools as a result of thermal shocks (Fig. 3b).

An incorrectly selected dose of the cooling lubricant may cause premature cracking of the tools and underfills of the tool`s cavity. It is the reason for an increased pressure value caused by the occurrence of an air pocket in this place, accelerating the formation of micro-cracks and the so-called Rebinder effect (Fig. 3c). The solution to this problem can be elimination of the human factor through the introduction of a precise dosing system and sequencing of the lubricant.

The development of automation has led to the emergence of newer solutions, which enable a more precise and accurate application of the lubricating material. At present, several types of technical solutions for the application of lubricating coatings can be distinguished [24]. Due to a high efficiency and a short time of lubricant application, automated solutions are realized on the basis of direct spraying. The latter is performed through a system of nozzles, which spray the lubricant onto the die during the change of the charge material. The construction of the nozzle depends on the applied lubricant and parameters, such as the spraying pressure and the lubricant's viscosity. An advantage of such a lubricating method is the easiness of directing the lubricant stream and the creation of a thin insulating layer on the surface of the tool. In addition, automated systems usually have a lubricant preparation station or a mixture thereof to maintain a uniform composition throughout the volume. A prior preparation of the lubricant reduces the possibility of deposition of the particulate matter in the pipes and nozzles, which results in a longer time between the service work performed on these devices. In order to obtain a uniform spray coating, the nozzles are positioned in such a way so as to reflect the curvature of the tool. The disadvantage of using classic nozzles is their high cost, size and the problem with the use of large-sized hammer dies.

The sole idea of automating the process of lubrication does not solve the analyzed problem. An example of this can be an automated method of lubrication on a crank-type press equipped with a transfer of forgings, as is the case, for example, of the process of a constant velocity joint body forging. In this process, the tools are lubricated by a specially constructed ejector with nozzles. Six nozzles evenly spaced around the periphery, every 60°, supply grease when the forging is ejected, when the ejector is in the upper position. The effect of this is the interior of the die being covered with a layer of grease with varying intensity, depending on the proximity of the nozzle, as well as an uneven wear (due to adhesion) on the perimeter of the die (Figure 4a).

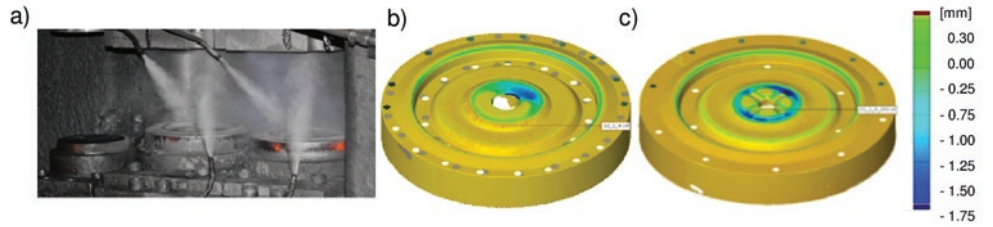


Fig. 2. a) Manual manner of lubricant supply, b) tool wear in the case of non-uniform lubrication and c) tool wear in the case of uniform lubrication

Fig. 3. Inappropriate lubrication: a) bad synchronization of the lubricating system with the forging process b) too intensive lubrication - no evaporation of the liquid from the lubricant mixture, c) crack in die inserts' corner – Rebinder effect

On the tool surface, one can observe a clear difference between the green and yellow areas (Fig. 4b) and the dark ones, reached by a smaller amount of the lubricating liquid. In the same process, in order to analyze the influence of the amount of the lubricant dosage on the changes in the temperature inside the tool (Fig. 5), the authors measured this parameter. A K-type thermocouple inserted into one of the dies was used for this purpose. The measurement showed that a two-fold increase in the amount of the lubricant causes a temperature drop at the surface of

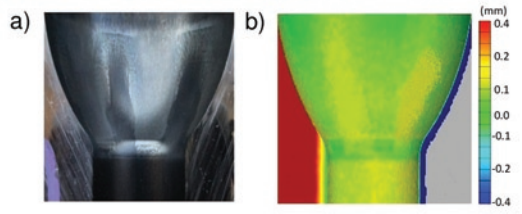


Fig. 4. Non-uniformly worn tool surface: a) digital image b) scanning results

the die of about 100° C. In addition, when analyzing the entire course, it can be observed how important it is to optimize and control the lubrication process for the entire life of the tool and not just a single forging operation.

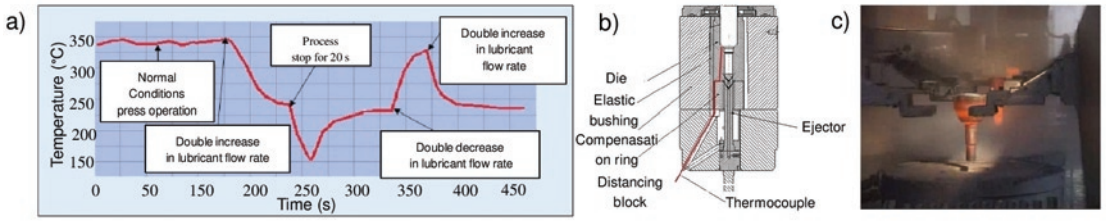


Fig. 5. a) A view of the die temperature change during forging with marked most frequently occurring process disturbances, b) the scheme of the thermocouple mounting into the tool, c) transfer of forgings on the transfer press

Similar studies regarding the temperature on the surface of forging tools during the forging cycle are presented in [8].

Currently, an increasingly common method of lubrication, especially in modern automated forging machines, is the use of manipulators' arms that accurately and repetitively insert the lubrication heads between the tools. These are high-performance systems that, because of their price and complexity, are more and more often used in cold forging [10]. The leading manufacturers of lubrication systems based on manipulators include: AED Automation, SMS Group, Renite, Spay. However, a currently used, yet much cheaper than the expensive manipulators, and at the same time a highly automated lubrication system designed for hot die forging is a device developed by Jerko Company (Fig. 6) [19]. This system has a special lubrication plate inserted between the upper and lower tools that applies a layer of grease. Such a solution will be used in Kuźnia Jawor (the Jawor Forge) as part of a research project.

In addition, the device enables a program selection of the nozzles to be active when feeding the lubricant, which further increases the flexibility of the system.

As shown above, despite the undeniable continuous technological development and the development of new technical solutions, the old, often unreliable and non-repetitive, lubrication devices are still used for most typical forging sets, which are additionally supported by the operators, responsible for the lubrication quality in the forging process. On the basis of the presented research and the analysis of the literature, it can be noticed that the key issue in the case of lubrication is primarily the method and direction of administration, as well as the amount (volume) and time of the lubricant-cooling agent dosage. These factors are of decisive importance, and if they are not properly selected, they significantly reduce the lifetime of the tools and may even cause premature damage to the tooling. This is also reflected in the reduced quality and accuracy of the dimensions and shapes of the forgings. The analysis makes it possible to conclude that the greatest influence on the wear of forging tools, omitting the aspect of the choice of the lubricant, is exerted by the precise setting of the volume of and the frequency of exposure to the lubricant-cooling dose. Somewhat less relevant is the direction of the lubricant feed to the surface of the dies. The optimal direction of spraying is the direction normal to the surface of the tool. However, due to the complicated tool shapes, it is in most cases impossible to implement. With an increasing complexity of the dosing system, i.e. extension of the spray head or automation system, the cost of the lubrication system increases significantly. The currently existing commercial solutions have been built based on the knowledge and experience of the companies creating these systems for selected industrial applications. In the literature, there is no information about the design and selection criteria for the proper lubrication system, especially for forging processes at elevated temperatures.

3. The concept of the proprietary lubrication system

In response to the presented state of knowledge, research results and demand, the authors have built, on the basis of their experience, a prototype lubrication and cooling system dedicated to the forging processes of yoke type forgings. The built-in dosing and distributing system is the first stage in the construction of a comprehensive, flexible lubrication system. In the next stage, it will be expanded by a manipulator and a nozzle system. This will enable further testing and determine the optimal method of proper spraying of the lubricant (Fig. 7). This is another solution of this type developed by the authors [15]. In the previously used lubricating apparatus in the process of forging a fork (yoke), the operator, by means of a "subjective" valve, regulated the dose of the lubricant and its exposure time through a lubricating nozzle. This resulted in a lack of lubrication stability, as well as a variability of the tribological conditions in the forging process. In comparison to the previous versions, the currently installed lubrication system has been partially modernized, and standardized components have been used, which enables easy servicing of the device in the industrial operation process as well as an easy access to the commonly available parts, which significantly reduces the cost of the device. The device has been made and tested on the production line at Kuźnia Jawor, which is the main recipi-

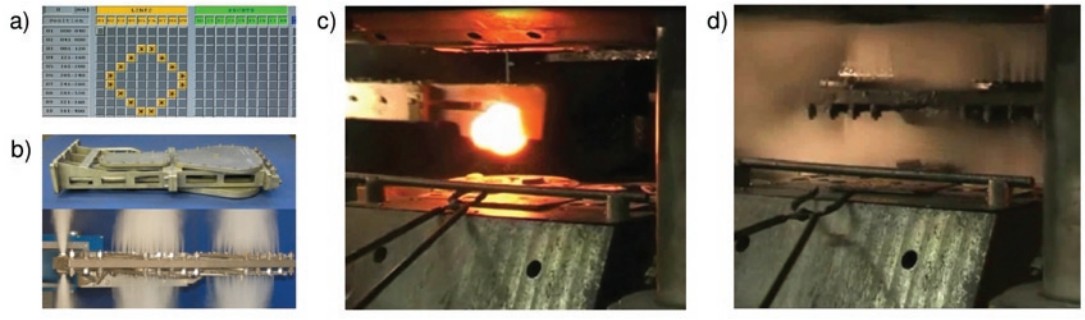


Fig. 6. The view of: a) the panel of the lubrication system with a system of active (selected) nozzles by Jerko[19], b) lubrication plate with nozzles, c) manipulator with gripper inserting a preform, d) entry of lubrication plate and lubrication

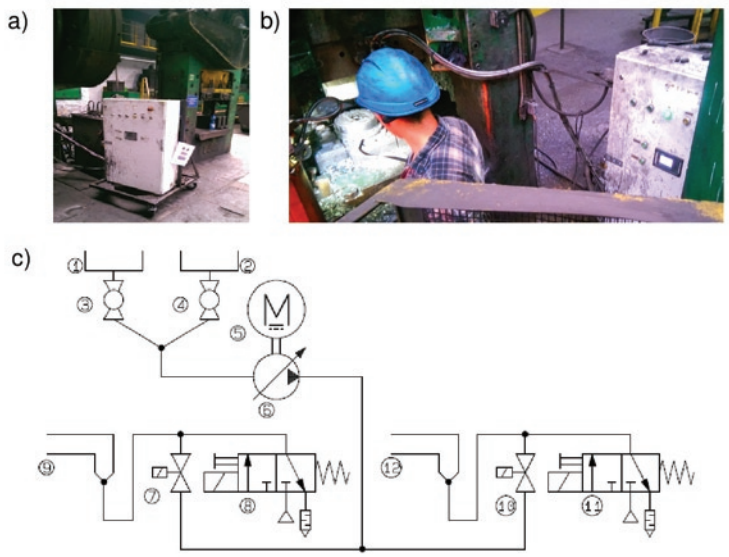


Fig. 7. a) A picture of the dosing and lubricating device allowing for adjusting the volume of the lubricant, b) the view of the connected and working device during the industrial forging process, c) diagram of the pneumatic system

ent of the system. The conducted tests made it possible to confirm the high efficiency of the hot forging process and allowed to improve it, which translates to high quality and lowered production costs. In the present solution, a particular attention was paid to the device's operating cycle (turn of sequences), as well as the number of doses and the direction of the lubricant application.

The described solution enables precise and repeatable metering of the dose to be dispensed, with the simultaneous possibility of varying the proportions for the subsequent nozzles, as well as the possibility of freely setting the sequence of the air blow and the application of the lubricant. For this purpose, a peristaltic pump (6) with a stepper motor (5) is used in the device shown in the diagram (Fig. 7c), which enables precise metering of the dispensed lubricant liquid (1).

The distribution of the dosed liquid to the upper (9) and lower (12) nozzles is carried out via two diaphragm valves (7 and 10) controlled pneumatically (8 and 11). After the quantity (volume) of the liquid is measured, the latter is distributed to each nozzle, i.e. the upper and lower nozzles. The developed solution makes it possible to push the lubricant to a special blow-out nozzle, developed by the authors, through flexible conduits using compressed air. The compressed air is additionally used to spraying the liquid, it is the blowing air. The order of feeding the lubricant-cooling medium, as well as the time of push and blow, are set by the operator using a simple HMI panel. A properly selected ejection time and amount of lubricant favor a proper spraying of the fluid, inhibiting the processes of graphite accumulation in the bends of the dies and the water remaining on their surface.

Fig. 8a shows a diagram of the device's operation and an exemplary work cycle diagram of the device. The regulation of the liquid phase content in the lubricant-cooling mixture consists in changing the operating time of the dosing pump. The device has an additional tank with clean water (designated as no. 2 in Fig. 7c) for cleaning, thus maintaining constant lubrication parameters. The choice between the container with the lubricating liquid and the one with the clean water is carried out with the use of ball valves (marked as no. 3 and 4 in Fig. 7c). The system is also equipped with an anti-sedimentation stirrer, which makes it possible to maintain the homogeneity of the graphite suspension in the water. The preliminary studies have shown a highly satisfactory operation of the device under industrial conditions. The next stage of development of the elaborated devices for the delivery of the lubricating-cooling agent will be the integration of the manipulator (Fig. 8b).

4. Comparison and analysis of the results of scanning performed on the applied forging tools by the newly developed and previously used lubrication system

Fig. 9 presents exemplary collective results obtained from the scanning of lower die inserts used in the preliminary forging of producing forgings type yoke with various numbers of forgings and under different tribological conditions (as a result of changing the lubrication system).

Before scanning, the inserts were cleaned of residual grease and scale. Next, the scanned digital images were processed to form a cloud of points and converted into a triangular mesh and then referred to the CAD model. The presented exemplary comparative results of the die insert scans (Fig. 9) transformed into a digital image indicate that, when the new lubricating-cooling device was used, enabling precise settings of the lubricant dose, a reduction in the tool wear was observed. The preliminary tests have shown that the use of 10ml of lubricant (a mixture of graphite with water in the ratio 1:20) during spraying for 2s provides the best results. In addition, it can also be observed that the provision of "better" lubrication conditions for the tool in the case of the insert (Fig. 9a) caused some displacements in the wear of selected tool zones compared to the tool with the old lubrication system. For example, the tool shown in Fig. 9b shows wear in places where the final parts of the forging arms are shaped, while for the tool shown in Fig. 9a (with the new lubricating system) no traces of wear are visible in this area (marked with a red ellipse). This can be explained by the fact that, as a result of the introduction of the newly developed lubrication system into the process, the tribological conditions changed, which caused changes in the occurrence of destructive mechanisms in selected areas of the tool.

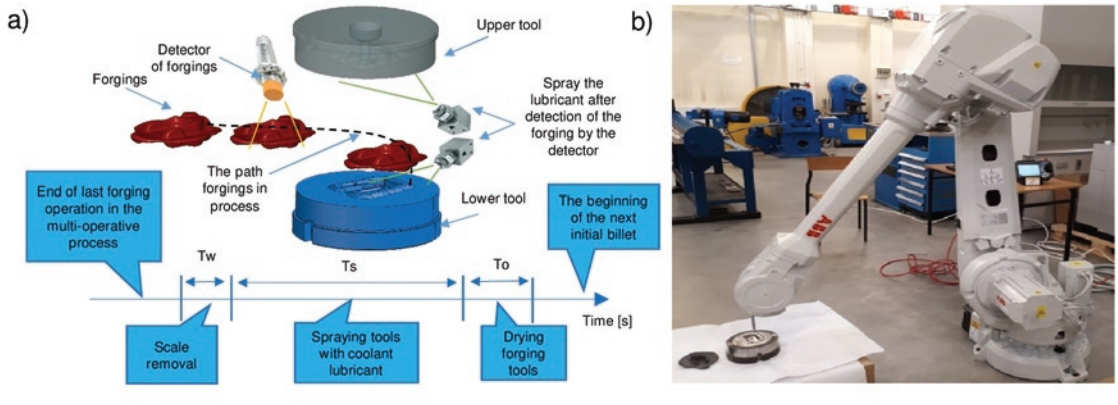


Fig. 8. a) The idea of a lubrication system operation with an exemplary work cycle diagram, b) using the manipulator arm for the targeted feeding of a lubricant-cooling agent

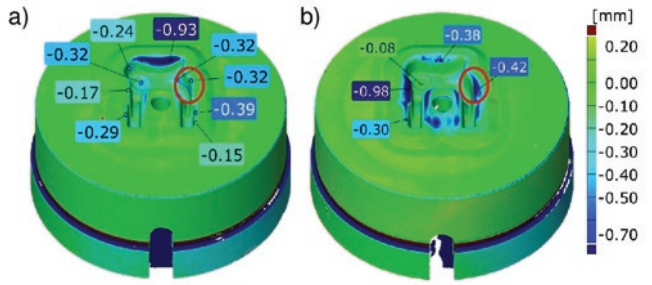


Fig. 9. Scan results of UNIMAX steel inserts used in the hot forging process: a) 16000 forgings with a new lubrication system, b) 16000 forgings with an old lubrication system

In order to determine the wear history of the selected die inserts (Fig. 9), the author's developed and partially improved reverse scan method was used, proved by the industrial practice to be a useful measurement and research tool. Until now, this method has been applied successfully to axisymmetric objects, with relatively simple shapes and large values of material loss [14,15]. The developed method consists in performing a measurement, with the use of a scanner, of the progressive wear of the selected forging tool (Fig. 10a), demonstrated as its loss of material, and on the basis of the shape changes of the forgings cyclically taken from the forging process, in the form of material increment (Fig. 10b). In order to reproduce the course of tool wear, the method involves scanning of the forgings selected from the production series (every 1000 item) from the total maximum number of forgings produced for each of the selected die inserts. To determine the graph describing the dependence of the volume consumption in

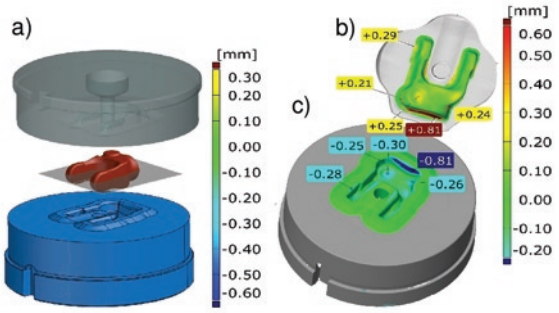


Fig. 10. The idea of the inverse method: a) set of forging tools for the second operation, b) comparison of the last forging results and c) matrix insert after 16000 forgings, in the form of changing the shape of the Surface

relation to the number of forgings made during the forging process, it is necessary to calculate the change in the measurement volume of the subsequent forgings to the result of 100 nominal forgings, which were leveled with the use of the reference surface.

In this way, it is possible to develop graphs showing the volume changes of the forgings cyclically collected during the process and then scanned. The measurement of the volume changes of the last produced forging for a given die insert is verified with the measurement of the tool volume at the end of the work. Fig. 11 presents an analysis of the volume changes carried out in order to determine the wear history of the selected lower die inserts for various lubrication conditions.

The proposed approach using the reverse 3D scanning method enables a global, but also thorough, description of the material loss, through observations of the volume changes during the forging process, as well as a quick, reliable and practical evaluation of the current state of the forging tools. In addition, the obtained results indicate that, on the basis of the course and shape (trend) of the consumption curves, you can predict the size of the material loss, and consequently control the amount of "tool opening" (flash thickness), and even determine the moment when the tool is withdrawn from further production due to its exceeding the dimensional tolerances.

Analyzing the determined wave forms of consumption curves in the aspect of the impact of the new lubrication system, it can be concluded that its introduction to the process provides more repeatable tribological conditions and reduced material consumption, especially in the final period of its exploitation. This leads to more advanced tests regarding both the optimization of the lubricant dose and the direction of its delivery, frequency and additional parameters related to the lubrication process.

5. Conclusions

The study has presented an analysis of the effect of the cooling and lubrication agents, their dosage amount as well as time and direction of application, and also the lubrication devices, on the wear of forging tools. As it has been demonstrated, the amount of the lubricant and the manner of its supply still constitute a current scientific problem. The performed analysis has shown that it is important to determine the optimal dosage of the lubricant as well as the way in which it is applied. Only through a complex long-term analysis of the given forging process it is possible to select the optimal tribological conditions ensuring stability and repeatability of the production process. That is why, at present, a big emphasis is put on the development of cooling and lubrication devices, as they are the ones which can ensure repeatable conditions of the tools operation, as opposed to the still commonly applied manual lubrication systems, affected by the subjective human factor. Although the die forging processes performed at elevated temperatures are similar, each of them requires an individual approach. And so, each process should be looked at separately and selected individually, which also refers to the parameters of its exposition and the lubrication technique. The cooling and lubrication system developed by the authors presented in the paper makes it possible to select and ensure the optimal tribological conditions of the process as

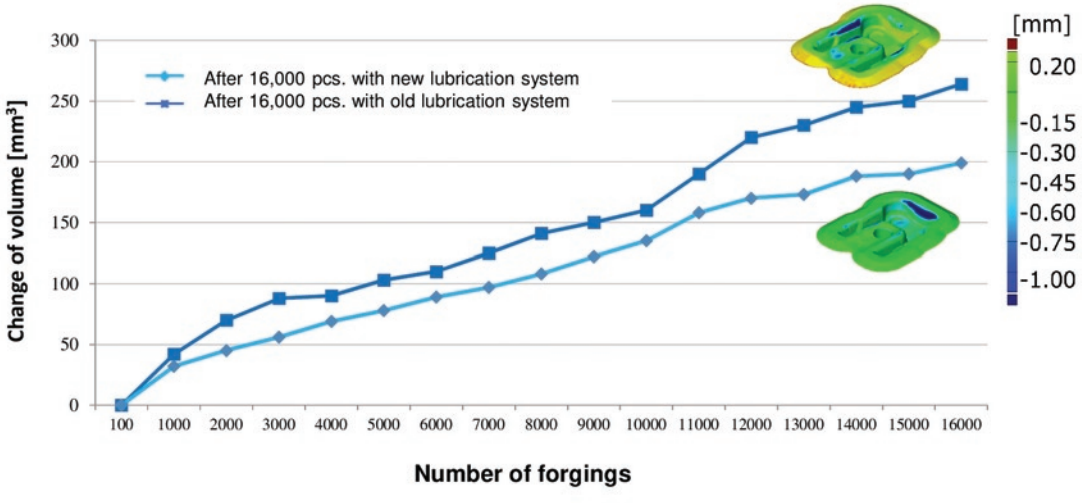


Fig. 11. Comparison of wear curves (history) determined on the basis of the original reverse scan method

well as to examine the effect of the amount and frequency of the lubricant dosage on the wear of the tool for the given process. The preliminary test results have demonstrated the usefulness of this device in the hot forging process of producing a fork type forging, in place of the currently applied, providing low stability, manually controlled lubricating devices. And so, it is justifiable to perform further, more advanced, research aimed at the selection of the optimal lubricant supply as well as other lubrication settings, which may also be connected with modernization of the elaborated device. Also, further investigations will be performed in relation to both the selection of the cooling and lubrication agent and its optimal temperature, in order to reduce the temperature gradient between the core and the working surface of the tools, causing thermal fatigue. The planned research also includes the use of a manipulator's arm for the examinations and tests of the selection of the proper direction and precision of the lubricant supply into the tool's impression.

Acknowledgements

This study was funded by National Centre for Research and Development, Poland (NCBiR); grant no. POIG.01.03.01-02-063/12.

References

1 Altan T, Ngaile G, Shen G. Cold and hot forging: fundamentals and applications. ASM International, Ohio 2005.
2. Altan T, Shirgaokar M. Advanced die materials and lubrication systems to reduce die wear in hot and warm forging. https://www.forging.org/uploaded/content/media/AltansPres.pdf
3. Anders P, Hogmark S, Bergström J. Simulation and evaluation of thermal fatigue cracking of hot work tool steels. International Journal of Fatigue 2004; (10): 1095-1107.
4. Archard J F. Contact and rubbing of flat surfaces. Journal of Applied Physics 1953; (24): 981-988, https://doi.org/10.1063/1.1721448.
5. Bay N. New lubricant systems for cold and warm forging – advantages and limitations. In Liewald, M.: Proceed. 12th Int. Cold Forging Congr., Stuttgart, Germany 2011; 1(8).
6. Bay N. New Tribo-systems for Cold Forming of Steel, Stainless Steel and Aluminium Alloys. Proceedings of 46th International Cold Forging Group (ICFG) Plenary Meeting 2013.
7. Berti G A, Monti M. Thermo-mechanical fatigue life assessment of hot forging die steel. Fatigue & Fracture of Engineering Materials & Structures 2005; 28 (11): 1025–1034, https://doi.org/10.1111/j.1460-2695.2005.00940.x.
8. Buchmayr B. Damage, Lifetime, and Repair of Forging Dies. BHM Berg- und Hüttenmännische Monatshefte 2017; 162 (3): 88–93, https://doi.org/10.1007/s00501-016-0566-3.
9. Choi Ch, Groseclose A, Altan T. Estimation of plastic deformation and abrasive wear in warm forging dies. Journal of Materials Processing Technology 2012; 212 (8): 1742–1752, https://doi.org/10.1016/j.jmatprotec.2012.03.023.
10. Colin S H. A review of automation in manufacturing illustrated by a case study on mixed-mode hot forging. Manufacturing Review 2014; (1)15, https://doi.org/10.1051/mfreview/2014012.
11. Daouben E E, et a., Effects of lubricant and lubrication parameters on friction during hot steel forging. International Journal of Material Forming 2008; 1: 1223–1226, https://doi.org/10.1007/s12289-008-0162-5.
12. Deacon R F, Goodman J K. Spreading Behavior of water based graphite Lubricants on Hot Die Surfaces. CIRP 2006; (55)1: 299-302.
13. Gronostajski Z, et al. The failure mechanisms of hot forging dies. Materials Science and Engineering, A Structural Materials: Properties, Microstructure and Processing 2016; 657: 147-160, https://doi.org/10.1016/j.msea.2016.01.030.
14. Gronostajski Z, Hawryluk M, et al. The application of the reverse 3D scanning method to evaluate the wear of forging tools divided on two selected areas. International Journal of Automotive Technology 2017; 18 (4): 653−662, https://doi.org/10.1007/s12239-017-0065-x.
15. Gronostajski Z, Hawryluk M, Kaszuba M, Ziemba J. Application of a measuring arm with an integrated laser scanner in the analysis of the shape changes of forging instrumentation during production. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18(2): 194–200, https://doi.org/10.17531/ein.2016.2.6.
16. Hawryluk M, et. al. Systems of supervision and analysis of industrial forging processes. Eksploatacja i Niezawodność - Maintenance and Reliability 2016; 18 (3): 315-324, https://doi.org/10.17531/ein.2016.3.1.
17. Hirschvogel M, Doelen H V. Some applications of cold and warm forging. Journal of Materials Processing Technology 1992; 35 (7): 343-356, https://doi.org/10.1016/0924-0136(92)90326-N.
18. https://www.aed-automation.com/en/products/solutions-forging.html, Aed Automation. Solutions forging. 2017. (accessed 17.11.02).
19. https://www.jerko-kempen.de, Jerko. 2017. (accessed 17.11.02).
20. Huskonen W D. Trends in Die Lubrication. Forging Magazine 2004; (10): 24-26.
21. Isogawa S, Kimura A, Tozawa Y. Proposal of an evaluating method on lubrication. CIRP Annals - Manufacturing Technology 1992; 41: 263-266, 10.1016/S0007-8506(07)61200-1.
22. Iwama T, Morimoto Y. Die life and lubrication in warm forging. Journal of Materials Processing Technology 1997; 71: 43-48, https://doi.org/10.1016/S0924-0136(97)00141-6.
23. Kima D H, Leeb H C, Kimc B M, Kimd K H. Estimation of die service life against plastic deformation and wear during hot forging processes. Journal of Materials Processing Technology 2005; 166: 372–380, https://doi.org/10.1016/j.jmatprotec.2004.07.103.
24. Kumar U. et al. Hot forging lubricants. International Journal of Mechanical Engineering and Robotics 2014; 3 (4): 155-163.
25. Lange K, Cser L, Geiger M, Kals J A G. Tool life and tool quality in bulk metal forming. Proceedings of the Institution of Mechanical Engineers. Part B: Journal of Engineering Manufacture November 1993; 207: 223-239, https://doi.org/10.1243/PIME_PROC_1993_207_085_02.
26. Lavtar L, Muhic T, Kugler G, Tercelj M. Analysis of the main types of damage on a pair of industrial dies for hot forging car steering mechanisms. Engineering Failure Analysis 2011; 18 (10): 1143-1152, https://doi.org/10.1016/j.engfailanal.2010.11.002.
27. Manji J. Die Lubricants, Forging 1994: 39–44.
28. Nagahama T, Enomae S. Cold- and warm-forging press developments and applications. Journal of Materials Processing Technology 1992; 35(3–4): 415-427, https://doi.org/10.1016/0924-0136(92)90331-L.

29. Ngaile G, Saiki H, Ruan L, Marumo Y. A tribo-testing method for high performance cold forging lubricants. Wear 2007; 262:684-692, https://doi.org/10.1016/j.wear.2006.08.009.
30. Persson A, Hogmarkb S, Bergstroma J. Thermal fatigue cracking of surface engineered hot work tool steels. Surface & Coatings Technology 2005; 191: 216– 227, https://doi.org/10.1016/j.surfcoat.2004.04.053.
31. Sagisaka Y, Ishibashi I, Nakamura T, Sasaoka E, Hayakawa K. Evaluation of Environmentally Friendly Lubricant for Aluminium Alloy Cold Forging. Steel Res. Int., Special Edition, Wiley-VHC Verlag, Weinheim, 2011; 245-248.
32. Schey J. Tribology in Metalworking: Lubrication. Friction, and Wear American Society for Metals, USA 1983.
33. Sheljaskow S. Current level of development of warm forging technology. Journal of Materials Processing Technology1994; 46 (7): 3-18, https://doi.org/10.1016/0924-0136(94)90099-X.
34. Sheljaskow S. Tool lubricating systems in warm forging. Journal of Materials Processing Technology 2001; 113 (1–3): 16-21, https://doi.org/10.1016/S0924-0136(01)00645-8.
35. Soltani M, Pola A, La Vecchia G.M, Modigell M. Numerical method for modelling spray quenching of cylindrical forgings. La Metallurgia Italiana 2015; (7)8 : 33-40.
36. Starlinga C, Brancob J. Thermal fatigue of hot work tool steel with hard coatings. Thin Solid Films 1997; 308(309): 436–442, https://doi.org/10.1016/S0040-6090(97)00600-7.
37. Taylan A, Gracious N, Gangshu S. Cold and hot forging fundamentals and application. ASM International. Asmmetalshandbook 2005; 14: 337-338.

Marek HAWRYLUK
Zbigniew GRONOSTAJSKI
Jacek ZIEMBA
Łukasz DWORZAK
Paweł JABŁOŃSKI
Department of Mechanical Engineering
Wroclaw University of Science and Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

Marcin RYCHLIK
Kuźnia Jawor S.A.
ul. Kuziennicza 4, 59-400 Jawor, Poland

E-mails: marek.hawryluk@pwr.edu.pl, zbigniew.gronostajski@pwr.edu.pl, jacek.ziemba@pwr.edu.pl, lukasz.dworzak@pwr.edu.pl, pawel.jablonski@pwr.edu.pl, marcinrychlik@kuznia.com.pl

Article citation info:
PUCHALSKI A, ŚLĘZAK M, KOMORSKA I, WIŚNIOWSKI P. Multifractal analysis vehicle's in-use speed profile for application in driving cycles. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 177–181, http://dx.doi.org/10.17531/ein.2018.2.02.

Andrzej PUCHALSKI
Marcin ŚLĘZAK
Iwona KOMORSKA
Piotr WIŚNIOWSKI

MULTIFRACTAL ANALYSIS VEHICLE'S IN-USE SPEED PROFILE FOR APPLICATION IN DRIVING CYCLES

MULTIFRAKTALNA ANALIZA EKSPLOATACYJNEGO PROFILU PRĘDKOŚCI POJAZDU W ZASTOSOWANIU DO TESTÓW JEZDNYCH*

Time signals recorded by the on-board diagnostic system (OBD), describing the manner of vehicle's movement in actual road conditions show non-stationarity and non-linearity, as well as statistical multiscalarity. In practice, it means that the analysis of registered time series requires modelling of non-linear phenomena. The aim of this study was to examine the nature of the vehicle speed profile in actual road conditions with the method of multifractal analysis. A number of studies indicates that the driving tests applied for many years have not been representative for the actual operating conditions of vehicles. For both the new Worldwide Harmonised Light duty Test Cycle (WLTC), a worldwide harmonised procedure of light vehicle testing, as well as in actual urban driving conditions along the measuring route, being subject to empirical research, confirmation of strong multifractal properties of the recorded vehicle speed time series have been obtained.

***Keywords**: multifractal analysis, driving cycles, actual road conditions.*

Sygnały czasowe rejestrowane przez system diagnostyki pokładowej OBD i opisujące sposób ruchu pojazdu w rzeczywistych warunkach drogowych, wykazują niestacjonarność i nieliniowość oraz statystyczną wieloskalowość. W praktyce oznacza to, że analiza zarejestrowanych szeregów czasowych wymaga modelowania zjawisk nieliniowych. Celem niniejszej pracy było zbadanie charakteru profilu prędkości pojazdów w rzeczywistych warunkach drogowych metodą analizy multifraktalnej. Szereg badań wskazuje, że stosowane przez wiele lat testy jezdne nie były reprezentatywne dla rzeczywistych warunków eksploatacyjnych pojazdów. Zarówno dla nowego cyklu jezdnego WLTC, światowej zharmonizowanej procedury badań pojazdów lekkich jak i w rzeczywistych warunkach drogowych jazdy miejskiej na trasie pomiarowej, będącej przedmiotem badań doświadczalnych uzyskano potwierdzenie silnych własności multifraktalnych rejestrowanych szeregów czasowych prędkości pojazdu.

***Słowa kluczowe**: analiza multifraktalna, testy jezdne, rzeczywiste warunki drogowe.*

1. Introduction

Tests of the performance of automotive drive systems find application in the phase of concept development (basic cognitive research), in the design phase (as part of prototype testing), in the production phase (during quality control), at the admission stage (type approval tests) and in the maintenance phase (diagnostic tests). This goal is attained through various types of examination at engine test stands, which differ with regard to the stipulated working conditions. This work undertakes to address the problem of analysing the in-use vehicle speed profile with the view to applying it to driving tests. The Vehicle Driving Cycle (VDC) is a time series of data representing vehicle speed, which is intended to reflect a vehicle's movement in real road conditions and be used in evaluation of a vehicle or engine in terms of its economy and emissions.

Many years of research have led to several hundred practical driving cycles in various countries and regions. Most of the current tests used to determine emissions of motor vehicles are developed in accordance with the principle of accurate simulation in the time domain. The most popular tests are ones like FTP-75 (Federal Test Procedure), NEDC (New European Driving Cycle) or JC08 (Japanese cycle). NEDC is used in Europe, the FTP 75 cycle is used in the United States and JC08 in Japan. NEDC is a synthetic cycle of the theoretical driving profile, while the other two come from actual usage data [3, 9].

More reliable results for fuel economy during the simulation of vehicle operating conditions at engine test stands are to be provided by the WLTC test (Worldwide Harmonised Light duty Test Cycle), which is a result of the world-wide harmonised light vehicle testing procedure. The cycle is based on actual driving profiles derived from statistical research [18].

Each methodology for creating driving tests requires an analysis of the nature of recorded in-use driving speed signals. Most of the actual time signals, including the time series recorded by the on-board diagnostic system (OBD) and describing the way the manner of vehicle movement in actual road conditions, show non-stationarity and non-linearity as well as statistical multiscalarity. In practice, it means that the analysis of recorded time series requires the use of non-linear analysis methods [1, 7]. The methods of multifractal formalism using local power measures, such as Holder signal regularity exponents or probabilistic indicators, are a good way of modelling the dynamics of such systems [5, 8].

The Holder singularity exponent, determined at each point where the function is defined, reflects the level of amplitude fluctuation in the vicinity of this point. The scaling factor of the probabilistic meas-

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

ure distribution function, based on the values of the signal amplitude, allows for segmentation with considerations for the level of entropy. The statistical distribution represented by the histogram of singularity exponents is a picture of the multifractality of the signal. There are two ways to determine the multifractal spectrum of a singularity exponent. The direct method relies on the approximation of the histogram for determined Holder exponents. The indirect method in which the fractal dimension is usually assumed to be the capacitive dimension of the curve being a graph of the signal considered, is based on the algorithm of detrended fluctuation analysis of time series [6].

The article presents the following issues. Chapter 2 signals the idea of the applied multifractal analysis method. The results of this analysis for the driving cycle of the harmonised light vehicle test procedure (WLTC) are presented in Chapter 3. The tests carried out in actual urban driving conditions are discussed in Chapter 4. Chapter 5 constitutes the concluding section.

2. Multifractal formalism and the regularity measure of a time series

Multifractal formalism has found a number of applications in the analysis of real signals, in particular in performance diagnostic tests of motor vehicles [4,11-17]. The conducted research indicates various procedures for identifying the multifractal nature of actual time series.

This chapter presents an approach based on the exponents of signal singularities and multifractal spectrum. The point exponent of Holder singularity function $f(x)$ in point x_0 denotes the number h defined as the supremum of all exponents satisfying, for a certain C> 0, the condition:

$$\left|f(x)-P_n(x-x_0)\right| \le C\left|x-x_0\right|^h \tag{1}$$

where $P_n(x-x_0)$ is a polynomial of $n<h$ degree. The relationship shows that the exponent $h>1$ describes the regularity of a function more accurately than its subsequent derivatives. If the signal representation is a time series $f_i, i=1,2,\ldots,N$, then:

$$f_{i+\Delta i}-f_i \sim \left|\Delta i\right|^{h(i)} \tag{2}$$

where $h(i)$ is the singularity exponent of time series at point i.

A set of fractal dimensions for each subset of time series elements f_i with the same Holder exponent h creates a multifractal spectrum of singularity. To determine the spectrum of Holder exponents, vehicle speed signals recorded by the OBD system we used a multifractal version of the detrending fluctuation analysis (MF-DFA) based on the elimination of the trend from the tested time series. The procedure leads to determining the measure in the form of a fluctuation function of row q exhibiting a power character:

$$F_q(s) \sim s^{H(q)} \tag{3}$$

where $H(q)$ is a generalised Hurst exponent, and the parameter q allows for decomposing the measure with regard to its value. Legendre's transform of generalised large-scale exponent $\tau(q)=qH(q)-1$ allows for obtaining a multifractal spectrum:

$$f(h)=qh-\tau(q) \tag{4}$$

where $h=\frac{d}{dq}\tau(q)$ is Holder singularity exponent.

The multifractal spectrum constitutes a histogram of exponents reflecting the level of signal amplitude fluctuation. Moreover, the parallelism of multifractal formalism and statistical thermodynamics indicates that Holder singularity exponent and the multifractal spectrum can be interpreted as the energy and entropy of the studied process respectively.

The description of the system's dynamic properties based on the multifractal spectrum of the time series is permitted by:

- the level $\Delta = h_{max} - h_{min}$ of multifractality determined by singularities with the largest and smallest time series fluctuation of the observed signal h_{min} and h_{max},
- singularity with the largest dimension, that is the most common time series singularity $\left\{h_0 : f(h_{0)} = \max f(h)\right\}$,
- the dimensional range of singularity subsets $\Delta f = f(h_{max}) - f(h_{min})$.

3. Simulation research of the WLTC test

A number of studies have confirmed that the driving tests used for many years are not representative for the actual operating conditions of vehicles. As a result, the emission and fuel consumption of vehicles are underestimated. Striving for a more dynamic, harmonised test cycle, a new WLTC (Worldwide harmonised Light duty Test Cycle) has been developed. The synthesis was made on the basis of data on traffic parameters in Europe, India, Japan, Korea and the USA, taking into account situations related to driving in urban, countryside and motorway traffic. Compared to NEDC, the test cycle is longer, significantly more dynamic and features a lot more acceleration and deceleration cycles, shorter stops and higher average and maximum speed values. Due to the introduced changes, the WLTC driving cycle will provide much more accurate conditions for calculating fuel consumption and exhaust emissions [10]. Its first application will apply to vehicle models introduced to the market for the first time since September 2017. The WLTC driving cycle has been divided into four parts, corresponding to different driving speeds: low, medium, high and extra high. If the maximum speed does not exceed 135 km/h, then part of extra high speed should be replaced with part of the low speed. The WLTC time series for 3b class cars, with the power to weight ratio of PWr> 34, is shown in Fig.1, and its basic parameters in Table 1. The drive dynamics in the case of these vehicles is determined by the driver's behaviour and the intensity of the road traffic, not by the car's technical parameters.

Singularity spectra for the studied WLTC driving test shown in Fig. 2, confirm its multifractal nature. The level of multifractality is: $\Delta = 7{,}92$, while the dimensional range of segments with the highest and lowest probability $\Delta f = 0{,}92$. The most common singularity exponents refer to those fragments of recorded time series which describe the greatest variability. In turn, the lowest probability of recording is shown by the drive periods of the highest regularity. Most of the points are concentrated around the dimensions corresponding to the singularities with the largest and smallest fluctuation of vehicles speed time series.

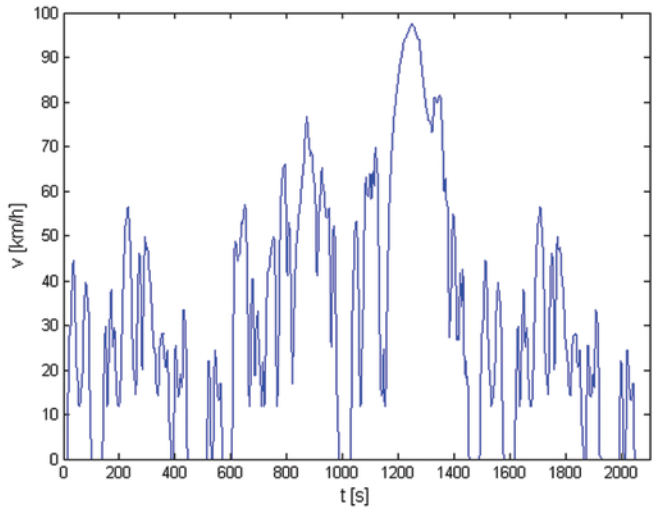


Fig. 1. WLTC driving cycle

Table 1. Basic parameters of the WLTC cycle for Class 3b vehicles

Phase	Duration	Stop duration	Distance	p_stop	v_max	a_min	a_max
	s	*s*	*m*		*km / h*	m / s^2	m / s^2
Low	589	156	3095	26.5 %	56.5	-1.47	1.47
Medium	433	48	4756	11.1%	76.6	-1.49	1.57
High	455	31	7162	6.6%	97.4	-1.49	1.58
Extra-High	323	7	8254	2.2%	131.3	-1.21	1.03
Total	1800	242	23 266				

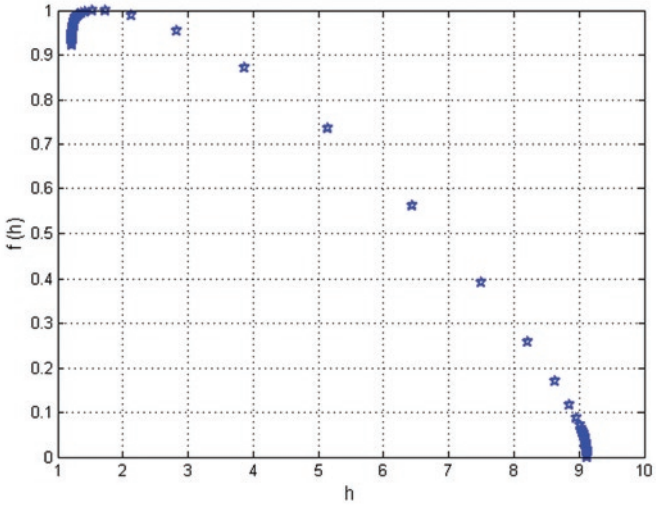


Fig. 2. The singularity spectrum for the WLTC driving test

4. Research and analysis of empirical data

The article presents the results of vehicle traffic tests in actual road conditions, represented by urban driving in a large agglomeration, between 9 am and 1 pm for five working days. The conducted analysis was based on vehicle's speed time series with a sampling period equal to 1s.

Fig. 3. CarChip Pro Recorder in a vehicle's OBD slot

The applied measuring device, shown in Fig. 3 is capable of sampling the speed once every second. The range of speeds available for recording is 0 -255 km/h, with a resolution of 1 km/h while the maximum recordable research distance is 16,000 km. The device has the ability to record a number of other operating parameters, such as engine speed, vehicle acceleration, pressure in the inlet manifold, and throttle position and allows communication with the engine governing unit using the following communication protocols: J1850-41,6, J1850-10,4, ISO9141, KWP2000 (ISO 14230), CAN (Control Area Network ISO 11898).

The measuring route consisted of a 12-kilometre-long road section and ran between Plac Wilsona and the Galeria Mokotów in Warsaw. The map presenting the measurement route, developed with the use of Google maps, is shown in Fig. 4.

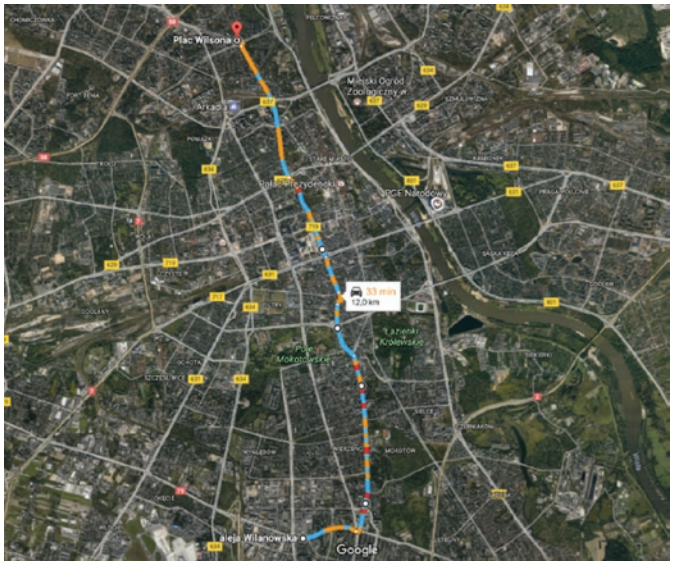


Fig. 4. Map showing the measurement route

Examples of speed change pattern recorded during one day for two-way trips are shown in Fig. 5. The conducted analysis was based on the vehicle's speed time series with a sampling period equal to 1s.

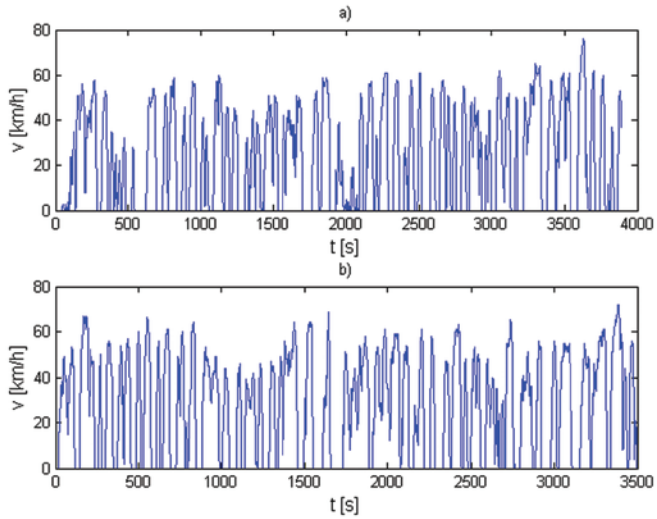


Fig. 5. Examples of vehicle speed change patterns

Multifractal spectra shown in Fig.6 - Fig.7 reflect the decomposition of the actual vehicle speed time series with reference to the abundance of fragments with a specific dynamics of variation.

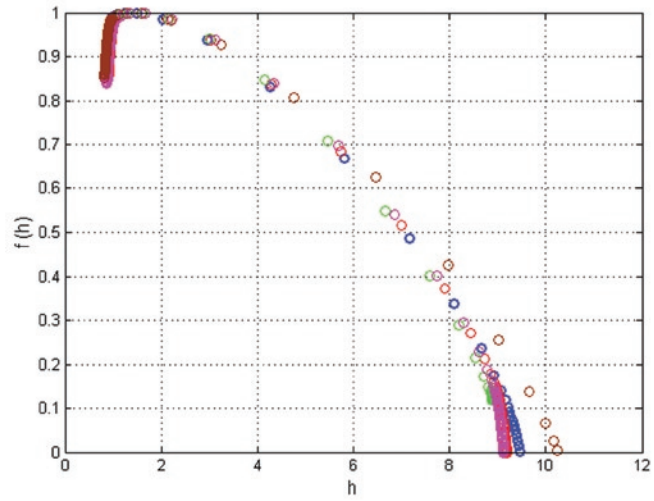


Fig. 6. Multifractal spectra of vehicle speed signals when driving from point A to B

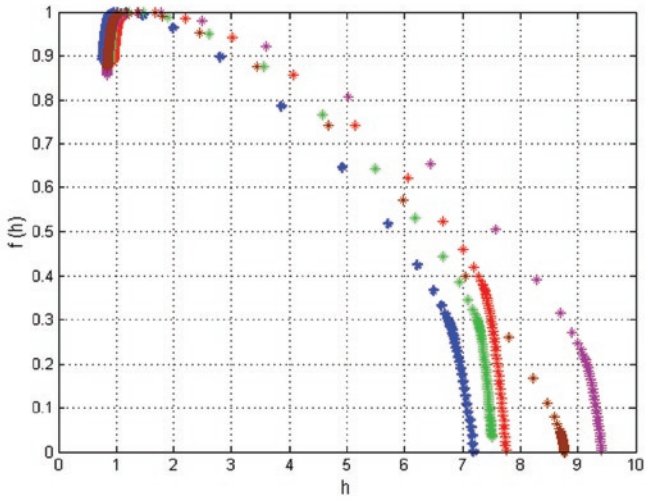


Fig. 7. Multifractal spectra of vehicle speed signals when driving from point B to A

All spectra have similar properties with regard to driving along the same route over consecutive five working days. They exhibit similar values of multifractality levels, singularities with the largest dimension and dimensional range of singularity subsets.

The segments of recorded speed signals exhibiting the largest fluctuations reach dimensions $f(h) > 0.8$. The lowest occurrence probability is shown by the regularity periods described by the singularity exponent $h > 7$. The multifractality level, Δ depending on the day of the week, reaches values in the range from 7 to 9.

5. Conclusions

The conducted analysis confirmed the multifractal nature of vehicle speed fluctuations both within WLTC driving cycle of the universal harmonised test procedure for light vehicles as well as real road conditions of urban driving along the measurement route, being subject of experimental research in Warsaw. Multifractal spectra are characterised by a similar shape and similar values of the level of multifractality. The asymmetric course of the spectra indicates the analogous properties for multifractals obtained as a result of the implementation of the generalised process of the binomial multiplicative cascade [2]. The obtained results indicate the possibility of using the multiplicative cascade process for the synthesis of real operating conditions and short-term forecasting of traffic conditions, which is a critical aspect in the development of intelligent transport systems. This step requires further research including driving in conditions outside urban centres and on motorways.

References

1. Batko W, Dąbrowski Z, Kiciński J. Nonlinear Effects In Technical Diagnostics. Radom, Poland: Institute for Sustainable Technologies, 2008.
2. Cheng Q. Generalized binomial multiplicative cascade processes and asymmetrical multifractal distributions. Nonlinear Processes in Geophysics 2014; 21: 477-487, https://doi.org/10.5194/npg-21-477-2014.
3. Chłopek Z. Synteza testów jezdnych zgodnie z kryteriami podobieństwa charakterystyk częstotliwościowych. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18(4): 572-577, https://doi.org/10.17531/ein.2016.4.12.
4. Dai M, Zhang Ch, Zhang D. Multifractal and singularity analysis of highway volume data. Physica A 2014; 407: 332-340, https://doi.org/10.1016/j.physa.2014.04.005.
5. Kantelhardt J W. Fractal and Multifractal Time Series. Mathematics of Complexity and Dynamical Systems Robert A Meyers (Ed.). New York: Springer 2011.
6. Kantelhardt J W, Zschiegner S A, Koscielny-Bunde E, Havlin S, Bunde A, Stanley HE. Multifractal detrended fluctuation analysis of nonstationary time series. Physica A 2002; 316: 87-114, https://doi.org/10.1016/S0378-4371(02)01383-3.
7. Kantz H, Schreiber T. Nonlinear time series analysis. Cambridge: University Press, 2004.
8. Loutridis S J. An algorithm for the characterization of time-series based on local regularity. Physica A 2007; 381: 383–398, https://doi.org/10.1016/j.physa.2007.03.012.

9. Merkisz J, Lijewski P, Fuc P, Siedlecki M, Ziółkowski A. Development of the methodology of exhaust emissions measurement under RDE (Real Driving Emissions) conditions for non-road mobile machinery (NRMM) vehicles. KONMOT 2016; Materials Science and Engineering 148: 1-11.
10. Mock P, Kühlwein J, Tietge U, Franco V, Bandivadekar A, German J. The WLTP: How a new test procedure for cars will affect fuel consumption values in the EU. Working paper 2014-9. International Council on Clean Transportation, 2014.
11. Puchalski A, Komorska I. Stable Distributions and Fractal Diagnostic Models of Vibration Signals of Rotating Systems. Advances in Condition Monitoring of Machinery in Non-Stationary Operations. Applied Condition Monitoring A.Timofiejczuk et al. (eds.) Springer International Publishing AG 2018; 9: 91-101, https://doi.org/10.1007/978-3-319-61927-9_9.
12. Puchalski A. Multiscale analysis of vibration signals in engine valve system. Journal of Vibroegineering 2015; 17(7): 3586-3593.
13. Puchalski A. Techniki budowy nieliniowych modeli diagnostycznych z wykorzystaniem szeregów czasowych drgań mechanicznych. Przegląd Mechaniczny 2016; 10: 33-36, https://doi.org/10.15199/148.2016.10.4.
14. Puchalski A, Komorska I. Multifractal Nature of Diesel Engine Rattle Noise in Vehicle. Archives of Acoustics 2017; 42(3): 469-474, https://doi.org/10.1515/aoa-2017-0049.
15. Shang P, Lu Y, Kamae S. The application of Holder exponent to traffic congestion warning. Physica A 2006; 370: 769–776, https://doi.org/10.1016/j.physa.2006.02.032.
16. Shang P, Lu Y, Kamae S. Detecting long-range correlation of traffic time series with multifractal detrended fluctuation analysis. Chaos, Solitions and Fractals 2008; 36: 82-90, https://doi.org/10.1016/j.chaos.2006.06.019.
17. Shang P, Shen J. Multi-fractal analysis of highway traffic data. Chinese Physics 2007; 16(2): 365-373, https://doi.org/10.1088/1009-1963/16/2/016.
18. Sileghem L, Bosteels D, May J, Favre C, Verhelst S. Analysis of vehicle emission measurements on the new WLTC, the NEDC and the CADC. Transportation Research Part D 2014; 32: 70–85, https://doi.org/10.1016/j.trd.2014.07.008.

Andrzej PUCHALSKI
Iwona KOMORSKA
University of Technology and Humanities in Radom
ul. Chrobrego 45, 26-600 Radom, Poland

Marcin ŚLĘZAK
Piotr WIŚNIOWSKI
Motor Transport Institute
ul. Jagiellońska 80, 03-301 Warszawa, Poland

E-mails: andrzej.puchalski@uthrad.pl, marcin.slezak@its.waw.pl, iwona.komorska@uthrad.pl, piotr.wisniowski@its.waw.pl

Article citation info:
LUTY W. Simulation-based analysis of the impact of vehicle mass on stopping distance. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 182–189, http://dx.doi.org/10.17531/ein.2018.2.03.

Witold LUTY

SIMULATION-BASED ANALYSIS OF THE IMPACT OF VEHICLE MASS ON STOPPING DISTANCE

SYMULACYJNA ANALIZA WPŁYWU MASY POJAZDU NA DROGĘ ZATRZYMANIA*

Results of experimental testing of motor truck tyres in dynamic braking conditions have been presented. With the measurement results being used as an example, higher normal wheel loads have been shown to result in longer time of rise in the longitudinal tangential tyre reaction force and in lower values of both the peak and sliding tyre-road adhesion coefficient. The data presented include results of simulation of the process of emergency braking of a motor truck whose mass can vary within wide limits. It can be seen from these results that an increase in the vehicle mass may considerably lengthen the vehicle stopping distance in emergency braking conditions.

***Keywords**: motor vehicle safety, stopping distance, tyre testing.*

W pracy przedstawiono wyniki badań eksperymentalnych ogumienia pojazdu ciężarowego w warunkach dynamicznego hamowania. Na przykładzie wyników pomiaru pokazano, że zwiększenie obciążenia normalnego koła skutkuje wzrostem czasu narastania wzdłużnej reakcji stycznej oraz spadkiem wartości współczynnika przyczepności opony do podłoża (przylgowej oraz poślizgowej). Przedstawiono wyniki symulacji procesu hamowania awaryjnego pojazdu ciężarowego, którego masa zmienia się znacząco. Wyniki wykazały ze zwiększenie masy pojazdu może istotnie wydłużyć jego drogę zatrzymania w warunkach hamowania awaryjnego.

***Słowa kluczowe**: bezpieczeństwo samochodu, droga zatrzymania, badania ogumienia.*

1. Introduction

The tyre-road adhesion ("grip") may be decisive for the vehicle behaviour in the conditions of extreme braking or drive along a road bend close to the limiting tyre-road adhesion [13, 14, 15, 17, 3]. In the braking process, the vehicle stopping distance may be expressed as [2, 12, 16]:

$$s_z = v_0\left(t_{rk} + \mathrm{t_{rs}} + \frac{t_n}{2}\right) + \frac{v_0^2}{2a_h} \tag{1}$$

where:

a_h – average braking deceleration
v_0 – initial vehicle velocity
t_{rk} – driver reaction time
t_{rs} – braking system response time
t_n – braking force/deceleration rise time

In the emergency braking conditions, the braking deceleration value a_h is limited by the adhesion force that can develop between the vehicle tyres and the road surface ("tyre-road adhesion force"). In the classic approach, the tyre-road adhesion force of each road wheel depends on the adhesion coefficient [1, 17, 3]. Hence, in the case of braking on a horizontal road and with an assumption made that the tyre-road adhesion coefficient μ_2 is constant (at wheel lockup), equation (1) takes the following form [2]:

$$s_z = v_0\left(\mathrm{t_r} + \frac{t_n}{2}\right) + \frac{{v_0}^2}{2\mu_2 g} \tag{2}$$

As it can be noticed, the vehicle mass is not present in this equation, which suggests that it does not influence the vehicle stopping distance at emergency braking. However, author's experience and literature data show that the tyre-road adhesion coefficient may vary with increasing normal wheel load [9, 8, 11] and a growth in vehicle mass may lengthen the emergency braking distance [18].

Changes in the mass (and weight) of passenger cars are in general rather small while the mass of present-day motor trucks may vary significantly, as the Maximum Authorized Mass (MAM), corresponding to the Gross Vehicle Weight (GVW), may be up to three times as big as the unladen mass.

Results of author's experimental research on vehicle tyres show that the normal wheel load value has a considerable impact on the course of changes in the values of the physical quantities that characterize the course of the braking process. This has been shown in Figure 1. The course of the whole process of dynamic wheel braking has been described in other publications [7, 8, 6]. It can be seen in example measurement results that in spite of applying a step signal U_h controlling the opening of the air brake control valve, the force F_{zh} that clamps the brake pads on the brake disc rises with a definite time delay and with a specific limited rate until it reaches its maximum value. A similar time delay and characteristic growth rate is observed for the longitudinal tyre slip s_x and longitudinal reaction force F_x (hereinafter referred to as "longitudinal reaction") transmitted by the tyre. A time history of this force has been shown in Figure 1b in the form of a unit force ($\mu_x = \frac{F_x}{F_z}$) vs time curve.

At the specific design of the disc brake calliper, the maximum value of the brake pad clamping force F_{zh}, which determines the wheel braking torque value M_h, is limited by the value of the air pressure applied to the brake actuator.

Therefore, a change in the normal wheel load value F_z should not be expected to cause changes in the time history and maximum value of the brake pad clamping force F_{zh} and, thus, of the wheel brak-

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

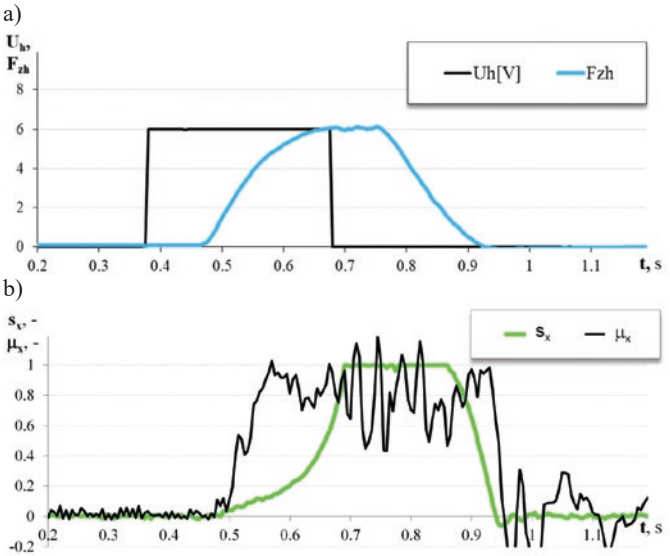


Fig. 1. Example set of results of measuring the physical quantities that characterize the process of dynamic braking of a medium-capacity motor truck wheel in laboratory conditions (normal wheel load F_z = 15 000 N, initial tyre rolling velocity v_0 = 60 km/h): a) voltage U_h of the brake valve control signal and brake pad clamping force F_{zh}, b) longitudinal tyre slip s_x and unit longitudinal reaction μ_x transmitted by the tyre

ing torque M_h. However, the wheel braking dynamics actually does change, as it can be seen in Figure 2.

The measurement results show that a growth in the normal wheel load F_z during dynamic braking of the wheel causes:

– lengthening of the time of drop in the angular wheel velocity until the wheel is locked up (Figures 2a, 2b);
– lengthening of the time of rise in the value of the longitudinal reaction F_x transmitted by the vehicle tyre until the value of the tyre-road adhesion force for the wheel locked up is reached (Figure 2b);
– decline in the peak (μ_1) and sliding (μ_2) tyre-road adhesion coefficient (Figure 2c).

Based on the presented results of laboratory tests of a wheel with a pneumatic tyre, and with reference to equation (2), a statement may be made that an increase in the vehicle mass directly causing a growth in the value of the normal load on each road wheel of the vehicle may result in an elongation of the vehicle stopping distance in the emergency braking process by:

– lengthening of the time of rise in the braking force up to a value corresponding to that of the tyre-road adhesion force;
– decline in the tyre-road adhesion coefficient.

These conclusions are important from the point of view of safety of vehicle motion and reconstruction of a road event during which emergency braking of a vehicle took place [15, 20, 19]. This problem chiefly applies to motor trucks, where the load mass may exceed the unladen vehicle mass.

The lengthening of the vehicle stopping distance due to an increase in the vehicle mass may be estimated by a simulation method. In the work described herein, simulation tests were planned and carried out which were aimed at presenting the impact of a growth in the vehicle mass and, thus, in the normal loads on vehicle wheels on the elongation of the vehicle stopping distance in an emergency braking process on the grounds of results of experimental tyre tests carried out in laboratory conditions.

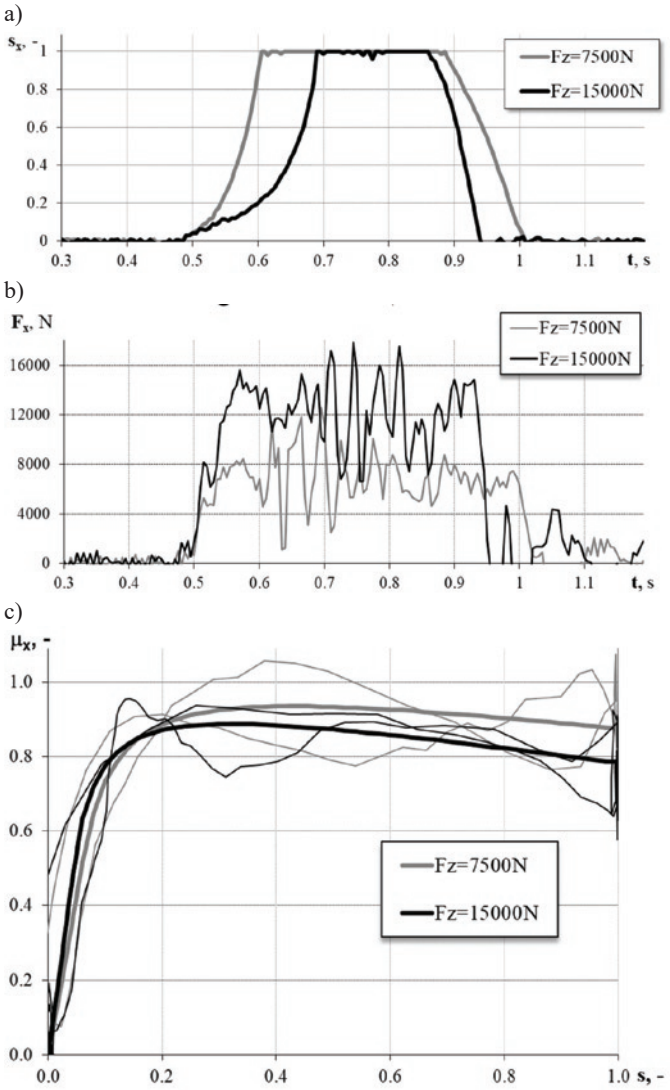


Fig. 2. Impact of the normal wheel load F_z on the course of dynamic wheel braking in laboratory conditions (medium-capacity motor truck wheel, v_0 = 60 km/h, road surface represented by a steel drum with a smooth surface): a) time histories of longitudinal tyre slip values s_x, b) time histories of the values of longitudinal reaction F_x transmitted by the tyre, c) comparison of wheel braking characteristics

2. Impact of normal wheel load on the process of growth in the value of the longitudinal tangential reaction transmitted by the tyre

To enable the execution of the simulation tests planned, a simplified description of the process of growth in the longitudinal tangential reaction transmitted by the tyre during dynamic braking had to be prepared and parametrized.

The process of growth in the longitudinal reaction F_x during dynamic braking of a vehicle wheel may be described in a simplified way by a linear relation, with the use of the following quantities (Figure 3):

– limiting longitudinal reaction value $F_{x,max}$, achieved and maintained during the wheel braking process;
– longitudinal reaction F_x rise time t_{nh}.

During the longitudinal reaction rise time t_{nh}, the angular wheel velocity ω is decreasing, which means a simultaneous growth in the longitudinal tyre slip s_x (Figure 3). For the purposes of this analysis, the limiting value $F_{x,max}$ of the longitudinal reaction F_x may be deter-

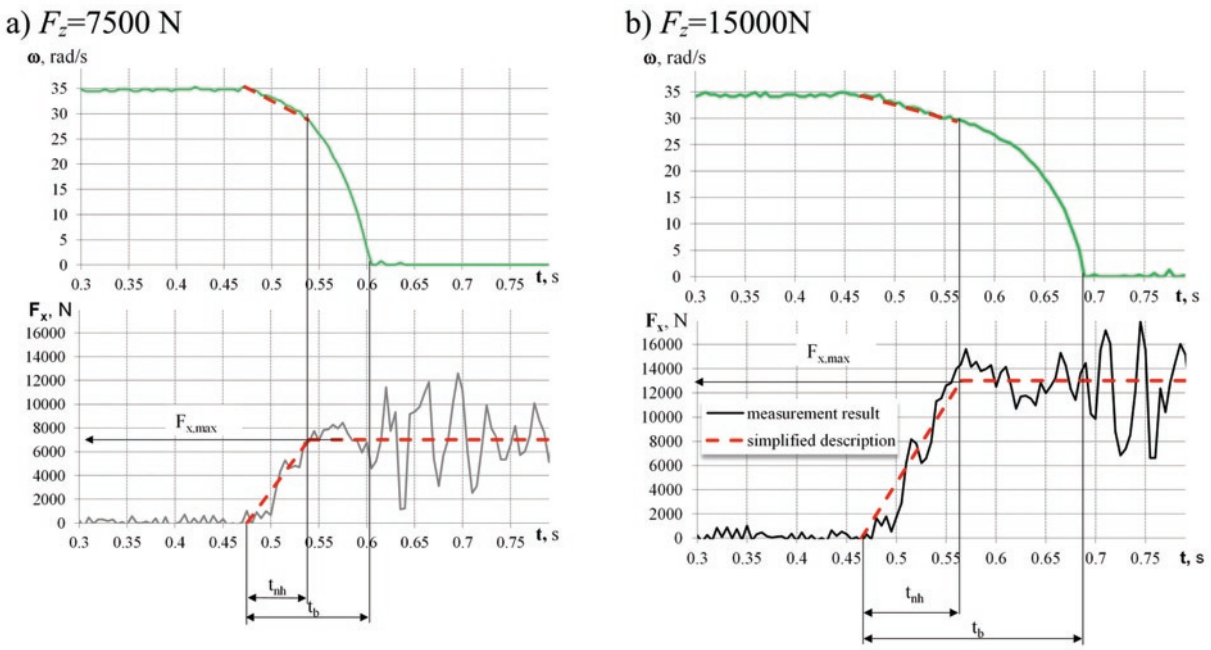


Fig. 3. Simplified description and parametrization of the dynamic wheel braking process

mined from the tyre-road adhesion coefficient. The results of author's experimental tests on a vehicle wheel with a tyre in braking conditions as presented in Figure 2 as well as the results of similar tests described in the literature [7, 8] show that the normal wheel load has an impact on the coefficient of adhesion of the individual wheel to the road surface. With a growth in the normal wheel load, both the peak and sliding tyre-road adhesion coefficient (μ_1 and μ_2, respectively) are declining. For the emergency braking, where vehicle wheels are locked up, an assumption may be made that the real value of the tyre-road adhesion coefficient is distributed around the sliding coefficient value μ_2 (Figure 3).

It can also be seen in Figure 3 that the limiting longitudinal reaction value $F_{x,max}$ was achieved in a time much shorter than the wheel lock-up time t_b. From the point of view of this analysis, it is important that in both cases under consideration, the longitudinal tangential reaction F_x reached the hypothetical limiting value $F_{x,max}$ corresponding to the sliding tyre-road adhesion force when the angular wheel velocity declined from the initial angular velocity ω_0 to a value of about $\omega = 4/5\ \omega_0$. Similar relative drops in the angular velocity occur when the wheel is braked from other initial velocity levels.

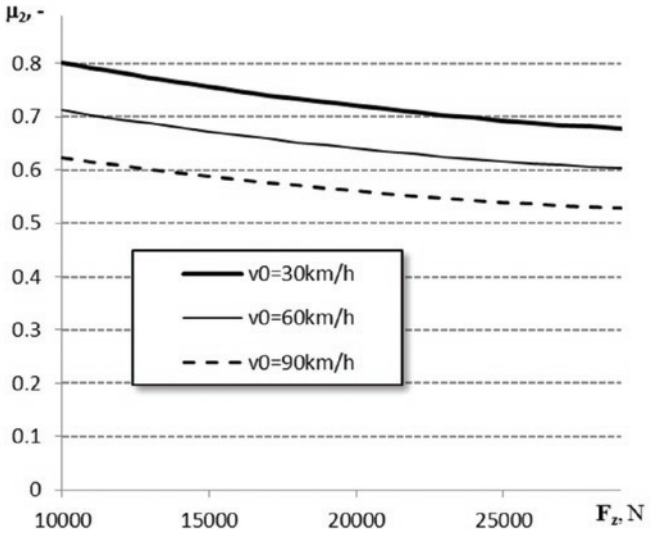


Fig. 4. Impact of the normal wheel load F_z on the sliding tyre-road adhesion coefficient values μ_2 (measurement results obtained in laboratory conditions, with the road surface being represented by a steel drum with a smooth surface)braking process

The impact of the normal wheel load on the tyre-road adhesion in the wheel lock-up condition, determined with taking into account the wheel rolling velocity, has been shown in Figure 4. The sliding tyre-road adhesion coefficient values μ_2, determined in a wide range of changes in the normal wheel load, may be directly used for estimating the vehicle stopping distance from equation (2).

Based on the conclusions drawn from an analysis of the measurement results, the following simplifying assumptions were adopted for the purposes of carrying out the simulation tests planned:

- in the dynamic braking process, the maximum braking torque M_h is applied to the wheel in a stepwise manner and its value is determined by the capacity of the wheel brake control mechanism,
- the braking torque rise time resulting from the inertia of the brake control mechanism ($t_{sh} \approx 0.2$ s) does not depend on the normal wheel load; this time was taken into account in the simulation process as a constant component of the braking force rise time $t_n = t_{sh} + t_{nh}$,
- in the period from $t = 0$ to $t = t_{nh}$, the rotational wheel motion is uniformly retarded, i.e. the angular wheel velocity ω linearly changes from the initial value of $\omega = \omega_0$ (defined by the wheel rolling velocity v_0 and the dynamic tyre radius r_d) to a value of $\omega \approx 4/5\ \omega_0$ (Figure 3),
- during the t_{nh} period, the longitudinal tangential reaction transmitted by the tyre is linearly rising from $F_x = 0$ to the limiting value Fx,max, following the formula $F_x(t) = \frac{F_{x,max}}{t_{nh}} t$ (Figure 3),
- the limiting value of the longitudinal tangential reaction is limited by the sliding tyre-road adhesion force defined by the formula $F_{x,max} = \mu_2 F_z$ (Figure 3)
- the value of the sliding tyre-road adhesion coefficient μ_2 depends on the normal wheel load F_z according to the relations shown in Figure 4.

The increase in the longitudinal reaction F_x rise time t_{nh} caused by a growth in the normal wheel load may be estimated on the grounds of an analysis of the dynamics of the rotational wheel motion during dynamic braking. A schematic diagram of the forces and torques acting on the braked wheel has been presented in Figure 5 [2].

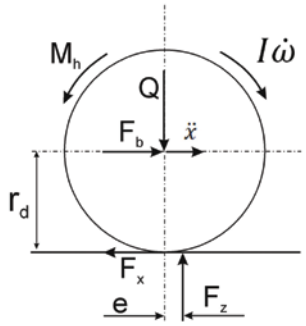


Fig. 5. Forces and torques acting on a vehicle wheel during the braking process.

Notation:

M_h – wheel braking torque
I – total moment of inertia of the wheel complete with the rotating elements connected to it
Q – wheel loading force, i.e. the part of the vehicle weight that is carried by the wheel
F_z – normal road reaction acting on the wheel
F_x – longitudinal road reaction acting on the wheel
ω – angular wheel velocity
$\ddot{x}$ – acceleration in the translational wheel motion
F_b – wheel pushing force, i.e. the part of the vehicle inertia force that acts on the wheel during the braking process

Based on the schematic diagram in Figure 5, the equation of vehicle wheel dynamics in the rotational motion has been formulated as follows:

$$I\dot{\omega} = -M_h - F_z e + F_x r_d \tag{3}$$

It may seem that the angular wheel deceleration during the braking process and, thus, the time of reaching the limiting value of the longitudinal reaction F_x are linearly related to the normal wheel load F_z. In the conditions of braking, however, the tyre is deformed in longitudinal direction by the longitudinal reaction F_x in accordance with the sense of the reaction force. Therefore, the above equation should be supplemented with an expression representing the impact of the longitudinal tyre deflection. The circumferential elasticity characteristics of present-day tyres are almost linear. Hence, for rough calculation purposes, the longitudinal tyre deflection values, starting from $u_x = 0$, may be expressed in a simplified form as [5, 10]:

$$u_x = \frac{F_x}{c_o} \tag{4}$$

where:

c_o – circumferential tyre stiffness.

With this supplement, the wheel dynamics equation will take the form:

$$I\dot{\omega} = -M_h - F_z\left(e - \frac{F_x}{c_o}\right) + F_x r_d \tag{5}$$

Additionally, the following supplementary assumptions may be made:

– For the wheel freely rolling, the shift e of the normal reaction is connected with the wheel rolling resistance coefficient by a relation:

$$e = f_t r_d \tag{6}$$

– The longitudinal tyre deflection u_x is linearly rising with the value of the longitudinal tangential reaction F_x in accordance with a relation:

$$u_x(t) = \frac{u_{xmax}}{t_{nh}} t = \frac{F_{x,max}}{c_o t_{nh}} t \tag{7}$$

In consequence, equation (5) will take the form:

$$I\frac{d\omega}{dt} = -M_h - F_z\left(f_t r_d - \frac{\mu_2 F_z}{c_o t_{nh}} t\right) + \frac{\mu_2 F_z r_d}{t_{nh}} t \tag{8}$$

By further transformations of this equation, with taking into account the simplifying assumptions adopted previously, the longitudinal reaction F_x rise time t_{nh} may be determined:

$$Id\omega = (-M_h - F_z f_t r_d + F_z \frac{\mu_2 F_z}{c_o t_{nh}} t + \frac{\mu_2 F_z r_d}{t_{nh}} t)dt \tag{9}$$

$$\int_{\omega_0}^{\frac{4\omega_0}{5}} Id\omega = \int_0^{t_{nh}} (-M_h - F_z f_t r_d + F_z^2 \frac{\mu_2}{c_o t_{nh}} t + F_z \frac{\mu_2 r_d}{t_{nh}} t)dt \tag{10}$$

$$I\frac{\omega_0}{5} = M_h t_{nh} + F_z f_t r_d t_{nh} - F_z{}^2 \frac{\mu_2}{2c_o} t_{nh} - F_z \frac{\mu_2 r_d}{2} t_{nh} \tag{11}$$

$$I\frac{\omega_0}{5} = t_{nh}\left(M_h + F_z f_t r_d - F_z{}^2 \frac{\mu_2}{2c_o} - F_z \frac{\mu_2 r_d}{2}\right) \tag{12}$$

$$t_{nh} = \frac{I\omega_0}{5\left(M_h + F_z f_t r_d - F_z{}^2 \frac{\mu_2}{2c_o} - F_z \frac{\mu_2 r_d}{2}\right)} \tag{13}$$

Results of the estimation of the time t_{nh} of rise in the longitudinal reaction F_x to the limiting value $F_{x,max}$ have been shown in Figure 6.

The value of the moment of inertia of a wheel with a tyre 275/70R22.5 complete with a hub and brake disc and the braking torque value were assumed on the grounds of test results given in other publications [7, 6].

Equation (13) shows a relation between the longitudinal reaction F_x rise time t_{nh} and many factors that characterize the conditions of wheel motion during the braking process. In the equation, direct dependence can be seen between the longitudinal reaction F_x rise time t_{nh} and the initial wheel rolling velocity v_0 (connected with the angular wheel velocity ω_0), although the impact of an increase in the initial wheel rolling velocity v_0 is partly compensated by the influence of a

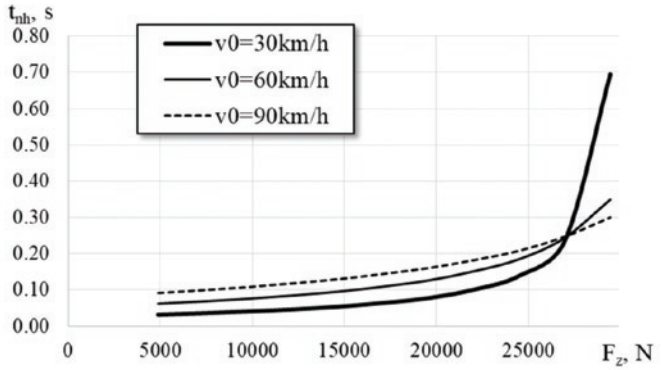


Fig. 6. Impact of the normal wheel load F_z and rolling velocity v_0 on the time t_{nh} of rise in the wheel braking force F_x to the limiting value $F_{x,max}$ (results obtained from simulations, for data typical of a medium-capacity motor truck)

decline in the tyre-road adhesion coefficient value μ_2. In effect, the simulation results have revealed that the time t_{nh} of rise in the wheel braking force F_x to the limiting value $F_{x,max}$ is nonlinearly growing with an increase in the normal wheel load F_z (Figure 6). The nonlinearity of this relation is particularly strong within the range of high values of the normal wheel load. Here, the impact of the circumferential flexibility of the tyre can be clearly seen as this flexibility, through the circumferential tyre stiffness c_o taken into account in equation (13), can significantly lengthen the estimated value of the time t_{nh}, especially at high values of the longitudinal reaction F_x, which are fostered by high values of the normal wheel load F_z.

The values of the longitudinal reaction F_x rise time t_{nh} are not very big. However, it has been shown that the time t_{nh} is considerably lengthened under the influence of growth in the normal wheel load F_z. Such an elongation causes the vehicle braking force rise time t_n to be lengthened, too. Thus, it may contribute to a lengthening of the vehicle stopping distance in accordance with the equation presented previously (2).

3. Evaluation of the impact of vehicle mass on the stopping distance

The normal load on vehicle wheels varies with changes in vehicle mass. Significant changes in vehicle mass may especially occur in the case of motor trucks, whose load capacity may even be twice as high as the unladen vehicle mass.

The results of experimental testing of motor truck tyres and calculation results were used for the simulation of the process of emergency braking of a motor truck with varying mass. Apart from the assumptions presented previously, the following simplifying assumptions had been adopted before a computing application was prepared:

- the vehicle moves on four wheels with comparable characteristics,
- the vehicle mass is uniformly distributed among individual road wheels,
- the driver starts the emergency braking process at the instant when a hazardous situation is noticed ($t = 0$),
- each vehicle wheel is subjected to a braking torque M_h of identical maximum value determined by the capacity of the wheel brake mechanism and its control system,
- the braking intensity is limited by the sliding tyre-road adhesion coefficient μ_2, whose value is determined at the beginning of the braking process for the initial vehicle velocity v_0 and depends on the normal load on each vehicle wheel, with the value of this load remaining unchanged during the braking process (in the simplified model adopted),
- the sliding tyre-road adhesion coefficient is identical for each wheel and is determined by vehicle weight and initial braking velocity,

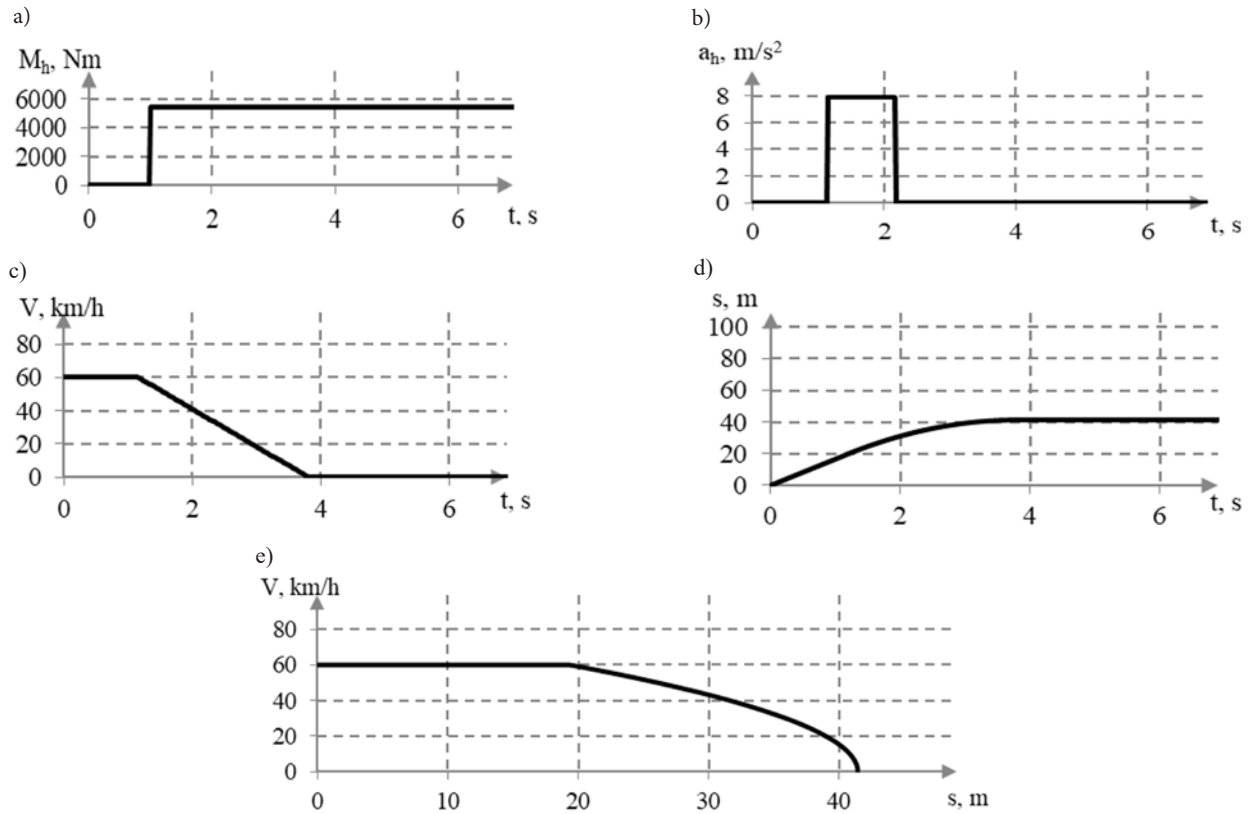


Fig. 7. Example set of results of calculating the physical quantities that characterize the process of emergency braking of a vehicle (v_0 = 60 km/h, m = 8 000 kg): a) braking torque acting on wheels, b) braking deceleration, c) vehicle velocity, d) distance travelled, d) distance travelled, e) Vehicle velocity vs distance travelled

– the dynamic tyre radius is identical for each wheel,
– any changes in the normal load on vehicle axles during the braking process were not taken into account,
– typical values of the driver reaction time and the braking system response time were adopted [20, 4]; however, these values as constants do not have any impact on the phenomena observed.

A spreadsheet making it possible to carry out the calculations planned was prepared. The simulation of emergency braking of a vehicle was based on equation (2) and on the modelling data described in Section 2. However, time histories of the vehicle velocity and distance travelled were obtained from iterative calculations, with determining (in predefined time intervals) successive values of the physical quantities that characterize the course of the braking process, including:

– braking deceleration a_h;
– vehicle velocity v;
– distance travelled s.

Pursuant to the assumptions adopted, each set of results was obtained for specific values of the sliding tyre-road adhesion coefficient μ_2, with the vehicle mass m and initial braking velocity v_0 being taken into account.

The calculations were carried out for the following options:

– initial vehicle velocity v_0 = 30, 60, and 90 km/h;
– vehicle mass m = 4 000, 8 000, and 12 000 kg, i.e. unladen, half-laden, and fully laden mass (MAM), respectively;
– road slope angle $\alpha = 0°$ (horizontal road).

An example set of calculation results has been presented in Figure 7.

From the point of view of the analysis carried out, the greatest importance is attached to the curve additionally plotted to represent the vehicle velocity v as a function of the distance travelled s, shown in Figure 7e.

Based on the example of the calculation results summarized in Figure 8, changes in the vehicle mass can be seen to have a definite impact on the quantities that characterize the vehicle braking process.

According to expectations, the calculation results showed that the raising of the vehicle mass, which means an increase in the normal load on each vehicle wheel, resulted in:

– time shift (delay) of the beginning of the braking phase (Figure 8a);
– reduction in the vehicle braking intensity and lengthening of the braking time (Figures 8a, 8b);
– lengthening of the vehicle stopping distance (Figure 8c).

However, the most conspicuous effects of an increase in the vehicle mass can be seen in Figure 8d. At a relatively low initial vehicle velocity (v_0 = 60 km/h), the raising of the vehicle mass from the unladen to the half-laden and fully laden (MAM) value caused

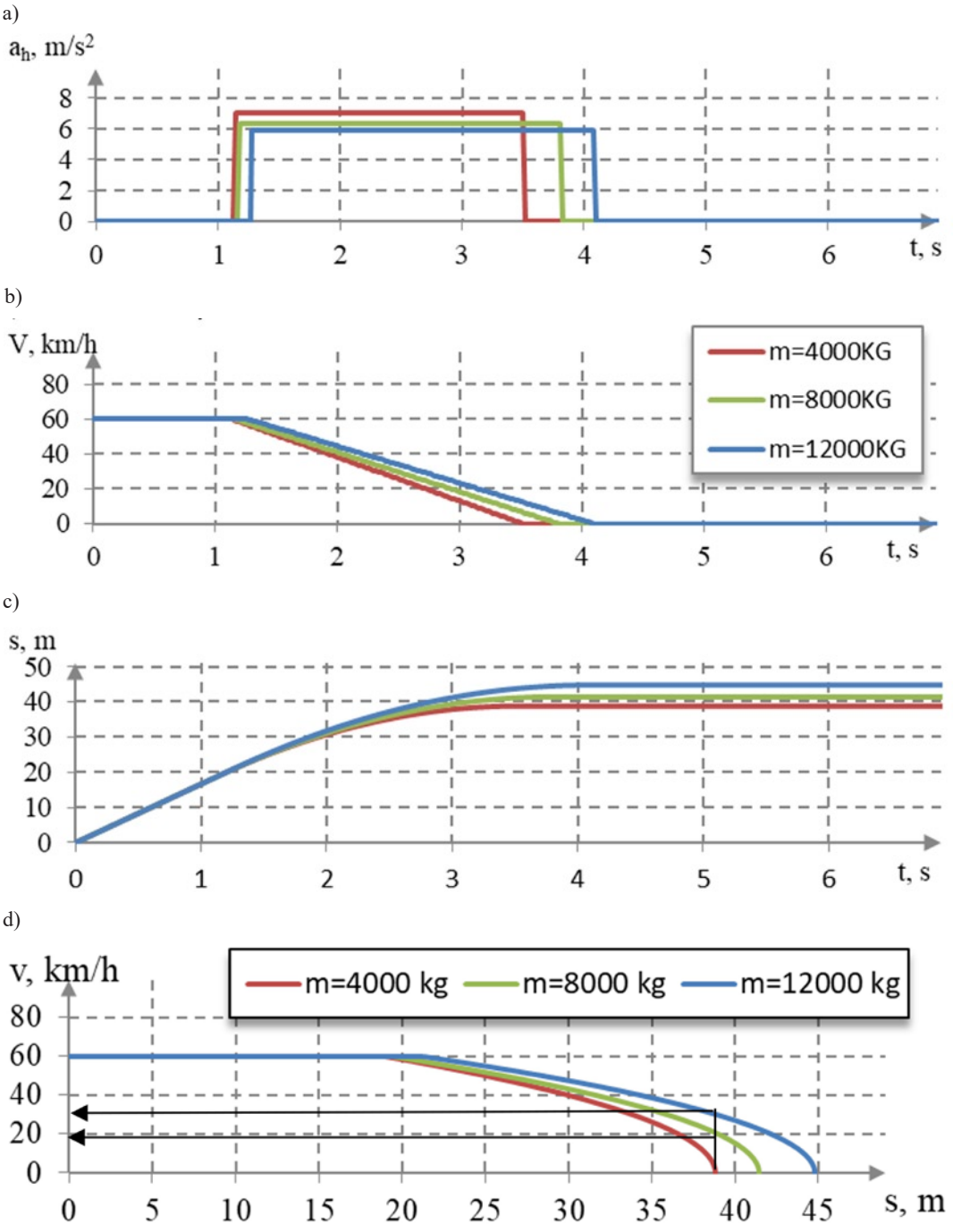


Fig. 8. Impact of vehicle mass on changes in the physical quantities that characterize the course of the emergency braking process (v_0 = 60 km/h): a) braking deceleration, b) vehicle velocity, c) distance travelled, d) vehicle velocity vs distance travelled

the vehicle stopping distance to be extended by about 4 m and 9 m, respectively.

These stopping distance elongation values are comparable with, respectively, the width of a typical pedestrian crossing and a half of the overall length of a typical tractor-semitrailer unit. On the other hand, conclusions of particular importance from the point of view of vehicle safety and reconstruction of a road event can be drawn from examining the results presented in Figure 8d. The calculation results have shown that at the place where the unladen vehicle (m = 4 000 kg) would stop, the velocity of the vehicle being half-laden (m = 8 000 kg) and fully laden (m = 12 000 kg) would be, approximately, over 20 km/h and over 30 km/h, respectively. In spite of moderate initial vehicle velocity, these residual velocity values are high enough for a possible collision between the vehicle and a pedestrian or another object to bring about very serious effects.

The tests revealed that the changes in the vehicle mass had an insignificant impact on the time of starting the braking process (Figure 9).

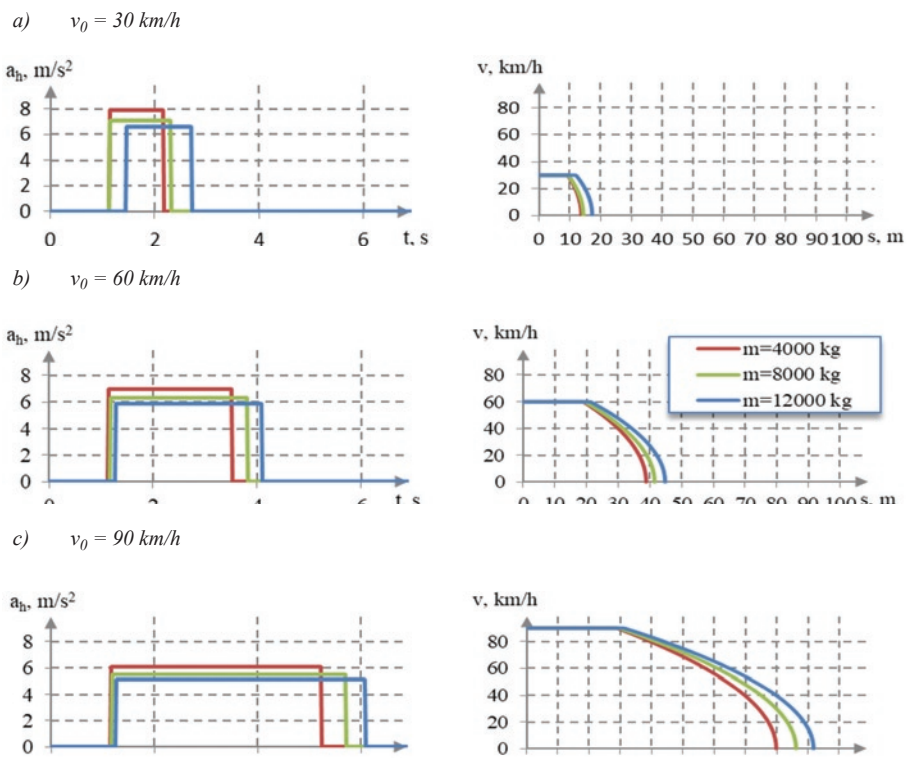


Fig. 9. Evaluation of the impact of vehicle mass m on the quantities that characterize the process of emergency braking of a vehicle from various initial braking velocities v_0

On the other hand, the increase in the vehicle mass caused an elongation of the time of rise in the braking force, especially at low values of the initial braking velocity v_0, according to the calculation results presented in Figure 6. Moreover, changes in the vehicle mass markedly affected the braking deceleration values and, in consequence, the braking and stopping distances achieved.

The impact of vehicle mass on the result of emergency braking in the conditions of various initial braking velocity v_0 has been summarized in Figure 10. At each initial braking velocity, an increase in the vehicle mass considerably lengthens the stopping distance. Simultaneously, it can be seen that the velocity of a fully laden vehicle (m = 12 000 kg) at the place where an unladen vehicle (m = 4 000 kg) would come to a halt may range from about **25** km/h to even **40** km/h, depending on the initial braking velocity.

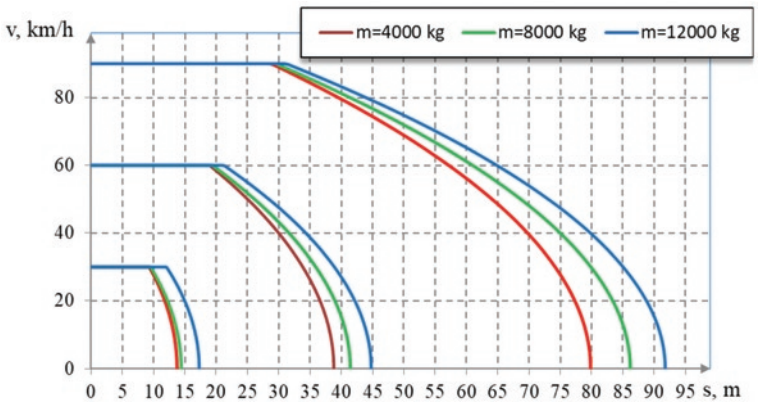


Fig. 10. Evaluation of the impact of vehicle mass m on the stopping distance at various initial braking velocities v_0

With the measurement results being used as an example, the following has been shown (Figure 11):

- the increase in the vehicle stopping distance caused by a growth in the vehicle mass is the highest at high initial braking velocities v_0,
- in relative terms, the raising of the vehicle mass from the unladen to the half-laden and fully laden (MAM) value may cause the vehicle stopping distance to be lengthened even by more than 20 %.

4. Closing conclusions

The simulation tests carried out have shown that the raising of the vehicle mass may considerably lengthen the emergency stopping distance of a vehicle in result of:

- delay in the start of the braking process;
- reduction in the braking intensity.

Moreover, it has been shown that the vehicle loaded with a cargo may still move with a considerable velocity at the instant when the unladen vehicle would have come to a halt. These conclusions are important from the point of view of safety of vehicle motion. Simultaneously, they show that significant changes in tyre properties, such as those indicated here, must be taken into account in the process of analysis and reconstruction of a road event.

The research work under consideration is worth continuing, in both its experimental and model simulation part. It has been shown that the phenomena of changes in the processes observed are rooted in the pneumatic tyre properties highlighted in the experiments. However, some simplifying assumptions were made in the simulation tests, which included a simplified model of friction between the pneumatic tyre when locked up and the road surface, with the adhesion coefficient value remaining constant during the whole braking process. It is presumed that the impact of the growth in the vehicle mass on the elongation of the vehicle stopping distance would be found stronger if the following factors were taken into account in the tests:

- real changes in the sliding tyre-road adhesion coefficient that occur with changes in the sliding velocity;
- reduction in the tyre-road adhesion coefficient during the significantly extended braking time.
- These issues may define the main directions for further research.

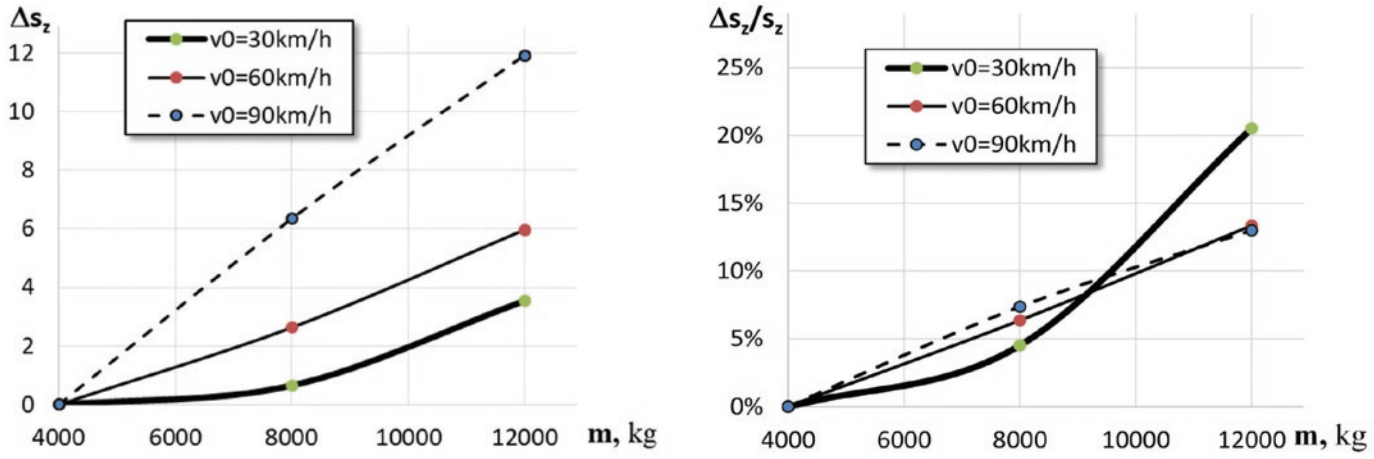


Fig. 11. Quantitative estimation of the impact of vehicle mass m on the increase in the stopping distance s_z in the conditions of emergency braking

References

1. Andrzejewski R. Dynamika pneumatycznego koła jezdnego (Dynamics of a wheel and pneumatic tyre assembly). WNT, Warszawa 2010.
2. Arczyński S. Mechanika ruchu samochodu (Mechanics of motion of a motor vehicle). WPW, Warszawa 1984.
3. Goudie D, Bowler J, Brown C, Heinrichs B et al. Tire Friction During Locked Wheel Braking. 2000. SAE Technical Paper 2000-01-1314.
4. Jurecki R, Jaśkiewicz M, Guzek M, Lozia Z, Zdanowicz P. Driver's reaction time under emergency braking a car – Research in a driving simulator. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 295–301.
5. Kulikowski K, Szpica D. Determination of directional stiffnesses of vehicles' tires under a static load operation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (1): 66–72.
6. Ławniczak S, Prochowski L. Analiza zmian prędkości kątowej koła samochodu ciężarowego podczas hamowania (Analysis of changes in the angular velocity of a motor truck wheel when braking). Zeszyty Instytutu Pojazdów / Politechnika Warszawska 2001; 1/40: 83-92.
7. Luty W, Prochowski L. Analiza procesu narastania siły hamowania koła samochodu ciężarowego (Analysis of the process of braking force rise in the motor truck wheel). Zeszyty Instytutu Pojazdów / Politechnika Warszawska 2001; 1(40): 9-24.
8. Luty W, Prochowski L. Modelowanie charakterystyk przyczepności ogumienia samochodów ciężarowych (Modelling of adhesion characteristics of motor truck tyres). Zeszyty Instytutu Pojazdów / Politechnika Warszawska 2002; 1(44): 37-47.
9. Luty W. Analiza właściwości nowych konstrukcji ogumienia kół jezdnych samochodu ciężarowego średniej ładowności (Analysis of the properties of medium-capacity motor truck tyres of new construction). In (joint publication): Analiza wpływu ogumienia nowych konstrukcji na bezpieczeństwo samochodu w ruchu krzywoliniowym (Analysis of the influence of tyres of new construction on the motor vehicle safety in curvilinear motion). Military University of Technology. Warszawa 2009; 7-18.
10. Luty W. Experimental research and analytical description of vehicle tires properties. Proceedings of the Institute of Vehicles / Warsaw University of Technology 2010; 1(77): 7-26.
11. Milliken W F, Milliken D L. Race Car Vehicle Dynamics, SAE International, Warrendale 1995.
12. Mitschke M. Teoria samochodu. Dynamika samochodu (Automobile theory. Dynamics of motor vehicles). Vol. 1: Napęd i hamowanie (Propulsion and braking). WKŁ, Warszawa 1987.
13. Parczewski K, Wnęk H. Make use of the friction coefficient during braking the vehicle. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (2): 176-180.
14. Parczewski K. Effect of tyre inflation pressure on the vehicle dynamics during braking manoeuvre. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 134–139.
15. Prochowski L, Unarski J, Wach W, Wicher J. Pojazdy samochodowe. Podstawy rekonstrukcji wypadków drogowych (Motor vehicles. Fundamentals of the reconstruction of road accidents). WKŁ, Warszawa 2008.
16. Prochowski L. Pojazdy samochodowe. Mechanika ruchu (Motor vehicles. Motion mechanics). WKŁ, Warszawa 2005.
17. Reed W, Keskin A. Vehicular Deceleration and Its Relationship to Friction. 1989. SAE Technical Paper 890736.
18. Sharizli A et al. Simulation and Analysis on the Effect of Gross Vehicle Weight on Braking Distance of Heavy Vehicle. Applied Mechanics and Materials 2014; 564: 77-82, https://doi.org/10.4028/www.scientific.net/AMM.564.77.
19. Warner C, Smith G, James M, Germane G. Friction Applications in Accident Reconstruction. 1983. SAE Technical Paper 830612.
20. Wiercinski J, Reza A et al. Wypadki drogowe. Vademecum biegłego sądowego (Road accidents. Forensic expert's vade mecum). Issue 2 (updated), Institute of Forensic Research Publishers (IES), Kraków 2010.

Witold LUTY
Warsaw University of Technology
Faculty of Transport
ul. Koszykowa 75, 00-662 Warszawa, Poland
E-mail: wluty@wt.pw.edu.pl

Article citation info:
BOŠNJAK S, GNJATOVIĆ N, MILENOVIĆ I. From 'a priori' to 'a posteriori' static stability of the slewing superstructure of a bucket wheel excavator. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 190–206, http://dx.doi.org/10.17531/ein.2018.2.04.

Srđan BOŠNJAK
Nebojša GNJATOVIĆ
Ivan MILENOVIĆ

FROM 'A PRIORI' TO 'A POSTERIORI' STATIC STABILITY OF THE SLEWING SUPERSTRUCTURE OF A BUCKET WHEEL EXCAVATOR

RÓWNOWAGA STATYCZNA NADWOZIA OBROTOWEGO KOPARKI KOŁOWEJ: OD MODELU "A PRIORI" DO MODELU "A POSTERIORI"

The complexity of the slewing superstructure (SS) balancing problem derives from the changeability of its geometric configuration, the complexity of working conditions as well as multiple limitations of the possible set of solutions. Having in mind the fact that the existing reference literature does not fully specify the procedure of static stability proof, the aforementioned procedure is presented in detail for the first time in this paper. A major contribution is represented in the classification of the slewing superstructure models into two groups which were named: the 'a priori' model (designed image of the SS) and the 'a posteriori' model (actual image of the SS). The fundamental stages of the 'a posteriori' model development method are presented in the paper. The transformation and validation of the calculation model 'a priori' to the calculation model 'a posteriori' was conducted on the basis of weighing results. A calculation of the basic parameters of the bucket wheel excavator (BWE) superstructure was conducted for both analyzed models by using the in-house developed software. The 'a posteriori' models provide a reliable calculation of the SS static stability and may be used not only for static stability proof of the machine as a whole, but also for load analysis of substructures and elements of BWE and related surface mining machines, such as spreaders. Besides that, the previously mentioned models are of extreme importance for a successful and reliable exploitation and maintenence of the machine since they present the basis for adjustment and control of limiting winch rope forces values, periodic control of mass and center of gravity position, as well as for a possible reconstruction which would be conducted in order to realize better customization of the machine versus operating conditions.

***Keywords**: bucket wheel excavator, slewing superstructure, static stability proof.*

Złożoność problematyki stabilizacji nadwozia obrotowego (slewing superstructure, SS) wynika ze zmienności jego konfiguracji geometrycznej, złożoności warunków pracy oraz wielu ograniczeń możliwego zbioru rozwiązań. Ponieważ istniejąca literatura nie opisuje w pełni procedury przeprowadzania dowodu na równowagę statyczną, niniejsza praca stanowi pierwszą próbę opracowania takiej procedury. Głównym wkładem niniejszego artykułu jest klasyfikacja modeli nadwozia obrotowego na dwie grupy: model "a priori" (zaprojektowany obraz SS) i model "a posteriori" (rzeczywisty obraz SS). W artykule przedstawiono podstawowe etapy metody opracowywania modelu "a posteriori". Walidacji modelu obliczeniowego "a priori" i jego transformacji do modelu obliczeniowego "a posteriori" dokonano na podstawie wyników ważenia. Dla obydwu analizowanych modeli wykonano obliczenia podstawowych parametrów nadwozia koparki kołowej przy użyciu oprogramowania własnego. Modele "a posteriori" zapewniają niezawodne obliczenia równowagi statycznej SS i mogą być stosowane nie tylko do sprawdzania równowagi statycznej maszyny jako całości, ale również do analizy obciążenia podzespołów i elementów koparki kołowej oraz powiązanych maszyn górniczych, takich jak zwałowarki. Poza tym wspomniane wcześniej modele mają ogromne znaczenie dla skutecznej i niezawodnej eksploatacji i konserwacji maszyn, ponieważ stanowią podstawę do regulacji i kontroli granicznych wartości sił liny wciągarki, okresowej kontroli masy i położenia środka ciężkości, jak również możliwej rekonstrukcji, którą przeprowadza się w celu lepszego dostosowania maszyny do warunków pracy.

***Słowa kluczowe**: koparka kołowa, nadwozie obrotowe, równowaga statyczna.*

1. Introduction

Identifying the weight and the centre of gravity (COG) position i.e. the basic parameters of static stability (BPSS) is of extreme importance during the development of bucket wheel excavators (BWEs) projects, as well as during their exploitation and maintenance [5]. This is the reason why validation of the BPSS calculation values is performed by weighing, conducted immediately after the first BWE erection and also after every reconstruction of the slewing superstructure (SS) [1,9,11,12].

The weight of the bucket wheel boom subsystem is dominantly influencing the intensity of the load of its hanging system i.e. winch rope forces [5]. Additionally, the SS COG position has a major influence on the load of the slew bearing [16,17]. According to the presented facts, it can be concluded that precise identification of BPSS enables: (a) a reliable adjustment of the winch rope forces limiting values, which present basic protection against the SS static stability loss, (b) an identification of the loads of the catch hooks, which prevent loss of the static stability in extreme load cases and (c) a determination of unevenness of slew bearing rolling bodies load distribution.

In reference literature in the field of bucket wheel excavators [9,15] – the problem of static stability is analyzed only in general. The influence of the BPSS on the loads of the bucket wheel boom stays is analyzed in [5], while the influence of counterweight (CW) and bucket wheel (BW) with drive mass on the dynamic response of the SS is analyzed in paper [4]. Dynamics of the large-scale load-carrying

structures is especially influenced by the slewing superstructure COG position [13,14].

The complexity of the SS static stability proof arises from the changeability of its geometric configuration and a relatively large number of partial loads combinations. In the literature which was available to the authors, only the calculation of the safety factor against overturning on the basis of project documentation i.e. before weighing (‘a priori’) is analyzed. Furthermore, such calculation is provided scarcely and in general. Also, the calculation algorithm with appropriate expressions is not given. The investigations, published in research papers [1,11,12], are entirely dedicated to the problems of experimental determination of the SS weight and its COG position. The static stability proof conducted on the basis of BPSS obtained by weighing (‘a posteriori’) is not even mentioned. With these facts in mind, the procedure of the SS static stability proof ‘a priori’ is presented, and the original procedure of static stability proof ‘a posteriori’ is developed and fully included in this paper. Research presented in this paper is the sequel of the research published in [5] thus the object of analysis is the same: BWE 1600 (Fig. 1 in [5]). Development of the SS ‘a posteriori’ models is of extreme importance if perceived in the light of the fact that BWEs are machines designed for perennial exploitation in harsh working conditions. During their exploitation life, which for the BWEs operating in Serbian open-pit mines exceeds 40 years, reconstructions due to technological [2] or structural [2,6] failures, as well as failures of vital mechanisms [10], or executed in order to realise revitalization/modernization of excavating units [3,15], are inevitable. By nature, the mentioned reconstructions are dominantly executed on the SS responsible for the realisation of the excavation process, and inevitably followed by an experimental control of BPSS and the eventual correction of the CW mass before putting the excavator back to exploitation. The use of ‘a posteriori’ static stability proof procedure enables the determination of actual values of the factor of safety against overturning for the reconstructed SS. Besides, regardless of the cause (check out or reconstruction), the SS weighing procedure causes direct (engagement of specialists and specialized equipment) and indirect material losses due to the system standstill. For example, three days are needed for the realisation of SS weighing of the BWE which is a vital part of the overburden excavation system on the Tamnava East field open-pit mine (Serbia), whereby indirect costs caused by the system standstill are equal to 9232 €/h [7]. Three weighing procedures were conducted prior to putting the BWE 1600 to exploitation [5]. The use of the ‘a posteriori’ static stability proof procedure would enable the correction of the CW mass on the basis of two weighing procedures, thus essentially reducing the indirect losses, while fully providing the stability of the machine.

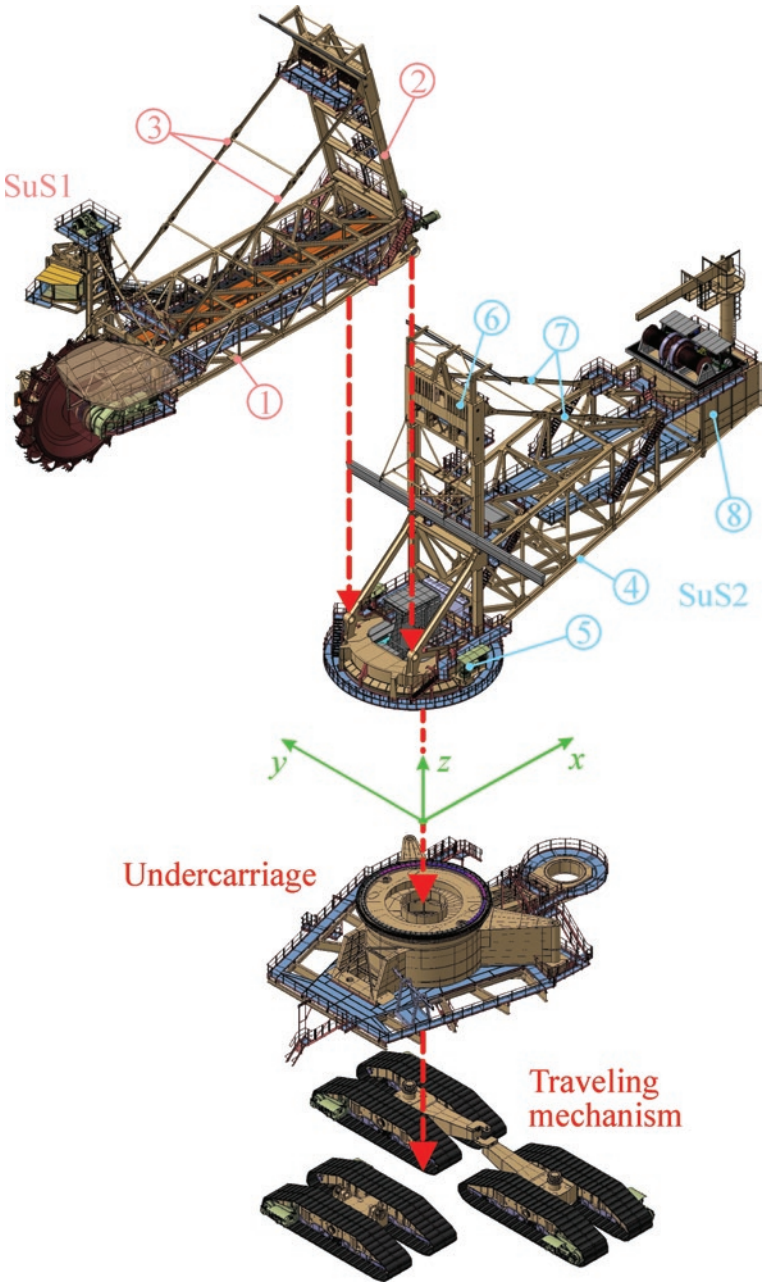


Fig. 1. Basic substructures of the BWE 1600: 1-bucket wheel boom; 2-mast 1; 3-bucket wheel boom stays; 4- counterweight boom; 5- slewing platform; 6-mast 2; 7-counterweight boom stays; 8- counterweight

The proposed static stability proof procedure ‘a posteriori’ has a wider significance because it can be successfully used when proving the static stability of spreaders and similar machines used in the systems for surface mining, as well as large-scale cranes.

2. Static stability of the slewing superstructure

The SS is leaning on the radiaxial slew bearing (RSB) which distributes the loads to the undercarriage and enables its slewing. For the static stability proof in the RSB plane, the following influences are accounted for:

- all of the loads acting upon substructure 1 (SuS1: bucket wheel boom (BWB)+mast 1 (M1)+BWB stays (BWBS), Fig. 1) and 2 (SuS2: counterweight boom (CWB) with slewing platform (SP)+mast 2 (M2)+CWB stays (CWBS), Fig. 1), as well as CW;
- ground reactions due to BW partial leaning;
- overload forces.

It is important to note that points of application of all the loads acting upon SuS1, in relation to the conditionally stationary coordinate system, are dependent of the BWB hoisting angle, which is not the case with loads acting upon SuS2.

The BWB hoisting mechanism of the BWE 1600 is not equipped with the system for continuous adjustment of winch rope forces extreme values. Their setting values (S_A: minimum - ‘’warning’’; S_{AA}: minimum - ‘’stop’’; S_Z: maximum - ‘’warning’’; S_{ZZ}: maximum - ‘’stop’’) remain constant on the complete domain of the BWB hoisting angle. The minimum and maximum rope force intensities ($S_{HZ3.1,min}$ and $S_{HZ3.1,max}$) obtained for load case (LC) HZ3.1 [8] (Tables 1 and

2) are representative for determining the winch rope forces setting values:

$$S_A = 0.94 S_{HZ3.1,min} \quad (1)$$

$$S_{AA} = 0.87 S_{HZ3.1,min} \quad (2)$$

$$S_Z = 1.08 S_{HZ3.1,max} \quad (3)$$

$$S_{ZZ} = 1.13 S_{HZ3.1,max} \quad (4)$$

Ground reactions caused by partial leaning of the BW and overload forces are determined for three characteristic LCs of the SuS1, Fig. 2:

- LC 1: dead load (G_{SuS1}=G_{BWB}+G_{M1}+G_{BWBS});
- LC 2: dead load (G_{SuS1}), weight of the material on belt conveyor 1 (F_1), incrustation of the BW (V_0) and conveyor 1 (V_1) and the digging force (U_F);
- LC 3: dead load (G_{SuS1}), weight of the material on belt conveyor 1 (F_1), incrustation of the BW (V_0) and conveyor 1 (V_1) and the digging force (U_L).

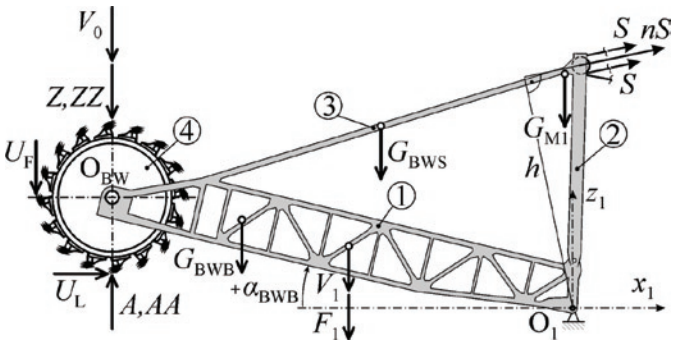


Fig. 2. The scheme for calculating the ground reaction due to partial leaning of the BW and overload forces: 1-BWB; 2-M1, 3-BWBS; 4-BW; S-winch rope force; n-number of rope legs; h-moment arm of pulley force; A, AA-ground reactions; G_{BWB}, G_{M1}, G_{BWBS}-weights of BWB, M1, BWBS; F_1-weight of material on belt conveyor 1; O_{BW}-center of the BW; U_F=U_L- nominal tangential component of the digging force; V_0 and V_1-weight of BW and conveyor 1 incrustation; Z, ZZ-overload forces; α_{BWB}-BWB hoisting angle

In the mentioned LCs, intensities of ground reactions (A, AA) are calculated on the basis of setting values of minimum winch rope forces (S_A and S_{AA}, expressions (1) and (2)), using the moment equation for y_1 axis, Fig. 2:

$$M_{y1,i} = S_A nh + \sum_L Lx_{1,L} - A_i x_{1,OBW} = 0, \quad (5a)$$

$$M_{y1,i} = S_{AA} nh + \sum_L Lx_{1,L} - AA_i x_{1,OBW} = 0, \quad (5b)$$

where: for LC 1, i=1, L=G_{SUS1}; for LC 2, i=2, L=G_{SUS1},F_1,V_0,V_1,U_F; for LC 3, i=3, L=G_{SUS1},F_1,V_0,V_1,U_L.

Expressing A_i and AA_i from equations (5a and 5b) respectively, while taking into consideration that according to [4] the calculation intensity of ground reaction is increased by 10%, the following equations are obtained:

$$A_i = 1.1 \frac{nh}{x_{1,OBW}} \left(S_A + \sum_L \frac{Lx_{1,L}}{nh} \right) = 1.1 \frac{nh}{x_{1,OBW}} \left(S_A - \sum_L S_L \right), \quad (6a)$$

$$AA_i = 1.1 \frac{nh}{x_{1,OBW}} \left(S_{AA} + \sum_L \frac{Lx_{1,L}}{nh} \right) = 1.1 \frac{nh}{x_{1,OBW}} \left(S_{AA} - \sum_L S_L \right), \quad (6b)$$

where S_L is the intensity of winch rope force induced by the influence of corresponding partial loading.

The overload forces are obtained by an analogous procedure, where as the basis for calculation, setting values of maximum winch rope forces (S_Z and S_{ZZ}, expressions (3) and (4)) are adopted:

$$Z_i = 1.1 \frac{nh}{x_{1,OBW}} \left(S_Z + \sum_L \frac{Lx_{1,L}}{nh} \right) = 1.1 \frac{nh}{x_{1,OBW}} \left(S_Z - \sum_L S_L \right) \quad (7a)$$

$$ZZ_i = 1.1 \frac{nh}{x_{1,OBW}} \left(S_{ZZ} + \sum_L \frac{Lx_{1,L}}{nh} \right) = 1.1 \frac{nh}{x_{1,OBW}} \left(S_{ZZ} - \sum_L S_L \right). \quad (7b)$$

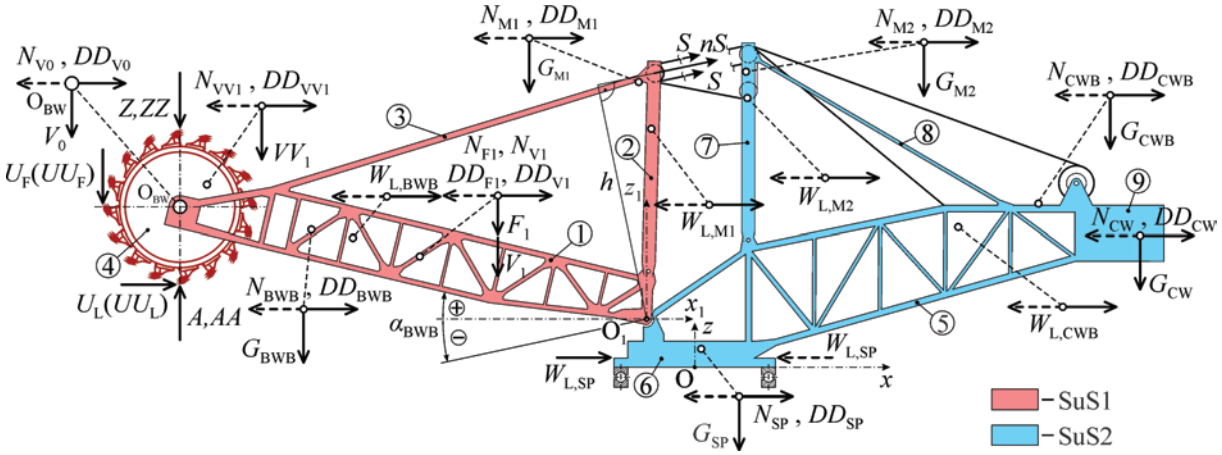


Fig. 3. The scheme of slewing superstructure loading: 1-BWB; 2-M1; 3-BWBS; 4-BW; 5-CWB; 6-SP; 7-M2; 8-CWBS; 9-CW; O_{BW}-center of the BW; S-winch rope force; n-number of rope legs; h-moment arm of the pulley force; G_{BWB}, G_{M1}, ..., G_{CW}-dead weights of substructures and CW; F_1-weight of material on belt conveyor 1; V_0 and V_1-weight of BW and conveyor 1 incrustation; VV_1-weight of the material in blocked BW chute; N_{BWB}, N_{M1}, ..., N_{CW}-longitudinal horizontal forces caused by normal inclination; NN_{BWB}, NN_{M1}, ..., NN_{CW}- longitudinal horizontal forces caused by transport inclination; $W_{L,BWB}$, $W_{L,M1}$, ..., $W_{L,CWBS}$-forces caused by longitudinal wind (in operation); $WW_{L,BWB}$, $WW_{L,M1}$, ..., $WW_{L,CWBS}$ -forces caused by longitudinal wind (out of operation); U_F=U_L-nominal tangential component of the digging force; UU_F=UU_L-maximum tangential component of the digging force; A, AA-ground reactions due to the partial leaning of the BW; Z, ZZ-overload forces; α_{BWB}-BWB hoisting angle; DD_{BWB}, DD_{M1}, ..., DD_{CW}-forces caused by earthquake

Table 1. Load combinations for BWS case

Loads	Load case[a]						
	H1.1	HZ3.1	HZ3.2	HZS4.4	HZS4.8	HZG5.3	HZG5.4
Group 1							
G	+	+	+	+	+	+	+
Group 2							
F_1	+	+	0	+	0	0	+
V_0	+	+	+	+	0	0	+
V_1	+	+	+	+	0	0	+
VV_1	0	0	0	+	0	0	0
W_L	0	$-x$[f]	0	$-x$	$-x$	$-x$	0
WW_L	0	0	$-x$	0	0	0	0
N[d]	$-x$	$-x$	$-x$	$-x$	$-x$	$-x$	$-x$
U_F	+	+	0	+	0	0	+
Z_1	0	0	0	0	+	0	0
ZZ_1	0	0	0	0	0	+	0
DD[e]	0	0	0	0	0	0	$-x$

[a]according to [8]; [b]sign "+" means that the load acts; [c]sign "0" means that the load does not act; [d]normal inclination (5%); [e]earthquake (11%); [f]load action caused by longitudinal wind, inclination and earthquake in "–x" direction (Fig. 3)

Table 2. Load combinations for CWS case

Loads	Load case[a]						
	H1.1	HZ3.1	HZ3.2	HZS4.1	HZS4.7	HZG5.2	HZG5.4
Group 1							
G	+	+	+	+	+	+	+
Group 2							
F_1	0	0	0	0	+	+	0
V_0	0	0	0	0	+	+	0
V_1	0	0	0	0	+	+	0
W_L	0	x[d]	0	0	x	x	0
WW_L	0	0	x	x	0	0	0
N	x	x	x	0	x	x	x
NN[c]	0	0	0	x	0	0	0
U	+/0[e] (U_L)	+/0 (U_L)	0	0	+ (U_F)	+ (U_F)	+/0 (U_L)
A_2	0	0	0	0	+	0	0
AA_2	0	0	0	0	0	+	0
DD	0	0	0	0	0	0	x

[a]according to [8]; [b]sign "+" means that the load acts; [c]sign "0" means that the load does not act; [c]transport inclination (10%); [d]load action caused by longitudinal wind, inclination and earthquake in "x" direction (Fig. 3); [e]influence of U_L is taken into consideration (+) for $z_{UL}>0$

When considering static stability, according to [9] all of the SS loads are divided into two groups. The first group (G1) consists of permanently acting forces of constant intensity (stabilizing forces – dead weight), while the second group (G2) consists of loads not acting permanently (overturning forces).

In general, depending on load conditions, and also on the geometric configuration of the SS (Fig. 3) violation of static stability is possible in two cases: (1) the loss of static stability on the bucket wheel side (case: BWS); (2) the loss of static stability on the counterweight side (case: CWS). In the first case the load combinations which give maximum winch rope forces are representative (Table 1) while in the second case the representative load combinations are those which cause minimum winch rope forces (Table 2).

[a]according to [8]; [b]sign ''+'' means that the load acts; [c]sign ''0'' means that the load does not act; [d]normal inclination (5%); [e]earthquake (11%); [f]load action caused by longitudinal wind, inclination and earthquake in ''–x'' direction (Fig. 3)

Static stability of the SS is obtained by its own dead weight:

$$G_{SS} = \sum_i G_i,\ \text{i=BWB, M1, BWBS, CWB, M2, CWBS, SP, CW,} \tag{8}$$

thus, its COG position:

$$\chi_{SS} = \sum_i G_i \chi_i,\ \chi = x, y, z,\ \text{i=BWB, M1, BWBS, CWB, M2, CWBS, SP, CW,} \tag{9}$$

related to the coordinate system *Oxyz* (whose applicate is coinciding with the SS slewing axis, Fig. 3) must be determined precisely. Abscissas and applicates of BWB, BWBS and M1 centers of gravity are dependent on the BWB hoisting angle (Fig. 2):

$$\chi_i = \chi_{1,i}(\alpha_{BWB}) + \chi_{O1},\ \chi=x,z,\ i=BWB, M1, BWBS, \quad (10)$$

which cause dependence of the complete SS COG position on the position of the BWB.

Positions of the points of application of the total vertical and longitudinal horizontal loading of the SS, caused by forces which belong to G2, are also dependent on the BWB hoisting angle. Besides, it is important to note that: (a) the direction of vertical loading belonging to G2 is constant, thus, it does not depend on the analyzed case of static stability loss (BWS or CWS); (b) the direction of longitudinal horizontal loading belonging to G2, with the exception of cutting force U_L, depends on the analyzed case of static stability loss; (c) all of the analyzed loads (Tables 1 and 2) are acting in planes parallel to plane *Oxz* and therefore do not have components along the direction of *y* axis (Fig. 3).

By reducing all forces acting upon the SS in point *O* (Tables 1 and 2, Fig. 3) the static system of the SS loading yields to the principal vector and principal moment with coordinates:

$$X = \sum_i X_i, \quad (11)$$

$$Z = \sum_i Z_i, \quad (12)$$

$$M_x = \sum_i Z_i y_i, \quad (13)$$

$$M_y = \sum_i (X_i z_i - Z_i x_i), \quad (14)$$

$$M_z = -\sum_i X_i y_i, \quad (15)$$

where X_i and Z_i are coordinates of partial loadings, while x_i, y_i and z_i are coordinates of the mentioned partial loadings points of application.

The $\vec{k}M_z$ component of the principal moment does not act upon the slew bearing, but upon the open gear transmission of the SS slewing drive. Therefore, the slew bearing is loaded by components $\vec{i}M_x$ and $\vec{j}M_y$ of the principal moment i.e. by the moment $\vec{M} = \vec{i}M_x + \vec{j}M_y$. The position of the moment $\vec{M}$ plane of action (Fig. 4) is determined by the angle:

$$\xi = \operatorname{arctg}\frac{M_y}{M_x}. \quad (16)$$

The acting plane of moment $\vec{M}$ is perpendicular to the *Oxy* plane, and their intersecting line (line *m*) forms with *x* axis the angle:

$$\xi_m = \xi - \frac{\pi}{2}, \text{ for } \xi > 0, \quad (17a)$$

$$\xi_m = \xi + \frac{\pi}{2}, \text{ for } \xi < 0. \quad (17b)$$

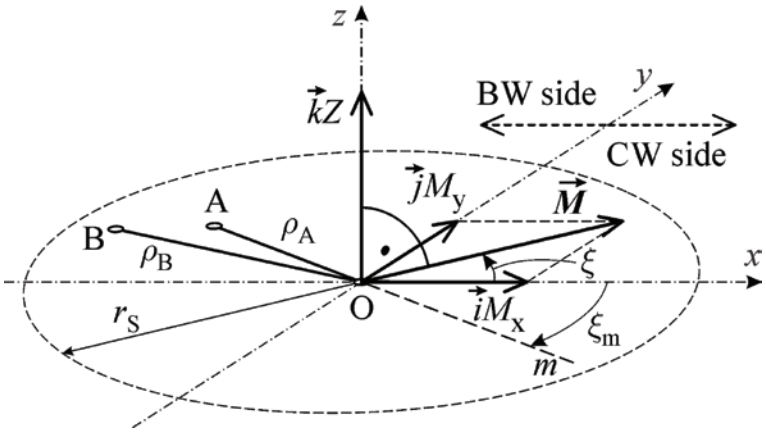


Fig. 4. The scheme for determining the distance of the total SS loading principal vector penetration point through referent plane of the RSB: Oxy-RSB referent plane; A-penetration point for the case of unfactored forces of both G1 and G2 groups; B-penetration point for the case of unfactored forces of group G1 and factored forces of group G2; $\vec{M}$ -moment acting upon RSB; $\vec{i}M_x$, $\vec{j}M_y$ -components of moment $\vec{M}$; $\vec{k}Z$ - component of the principal vector; ξ-angle between the moment $\vec{M}$ and the x axis; m-intersecting line between the moment $\vec{M}$ plane of action and the RSB referent plane; ξm-angle between line m and x axis; rS-radius of stability contour; ρA-distance of point A from the slewing axis; ρB-distance of point B from the slewing axis

The $\vec{k}Z$ component of the principal vector is perpendicular to moment $\vec{M}$ (Fig. 4), meaning that their action is equivalent to the action of force $\vec{k}Z$ with line of action passing through point *A* with coordinates:

$$x_A = -\frac{M_y}{Z}, \quad (18)$$

$$y_A = \frac{M_x}{Z}. \quad (19)$$

Point *A* belongs to the line *m* and is at the distance of:

$$\rho_A = \sqrt{x_A^2 + y_A^2} \quad (20)$$

from point *O*. Point *A*, determined in such a way, also represents the point of penetration of the total loading (unfactored forces of both G1 and G2 groups) of the SS principal vector line of action.

According to [8] contour of stability, i.e. the calculating contour of leaning in the plane of the slew bearing, is a circle with a radius of:

$$r_S = 0.95 r_{SB}, \quad (21)$$

with r_{SB} being the radius of the slew bearing rolling path.
The necessary condition of the SS static stability is:

$$\rho_{A,max} < r_S, \quad (22)$$

where $\rho_{A,max}$ is the maximum distance of point A from the SS slewing axis.

The minimum values of safety factors against overturning, prescribed by the standard [8] are dependent on the analyzed load case: for LCsH, $v_{H,min}$=1.5; for LCsHZ, $v_{HZ,min}$=1.33; for LCsHZS, $v_{HZS,min}$=1.2; for LCsHZG, $v_{HZG,min}$=1.1. When the forces belonging to G2 are factored with corresponding values of safety factors and the calculation is conducted according to expressions (11-20), coordinates of point B (point of penetration of the principal vector line of action in case of unfactored forces of group G1 and factored forces of group G2) and its distance from the SS slewing axis can be obtained. The sufficient condition of the static stability is:

$$\rho_{B,max} < r_S , \tag{23}$$

where $\rho_{B,max}$ is the maximum distance of point B from the SS slewing axis (Fig. 4).

The penetration point (C) of the overturning forces principal vector (unfactored forces of group G2) line of action through the referent plane of slew bearing, whose position is determined according to expressions (11-20), belongs to line n which intersects the stability contour in point T (Fig. 5). Tangent t on the stability contour in point T is the line of possible overturning.

Distance between points C and T is:

$$l_T = \rho_{C,max} - r_S . \tag{24}$$

The length of the line segment $\overline{ET}$ (Fig. 5) and the distance between points C_{SS} and F are determined according to expressions:

$$\overline{ET} = r_S \sin\xi_n + y_{SS} , \text{ for case BWS;} \tag{25a}$$

$$\overline{ET} = r_S \sin\xi_n - y_{SS} , \text{ for case CWS;} \tag{25b}$$

$$\overline{C_{SS}F} = x_{SS} + r_S \cos\xi_n + \overline{EF} = x_{SS} + r_S \cos\xi_n + \overline{ET}\,\mathrm{tg}\xi_n = \\ = x_{SS} + r_S \cos\xi_n + (r_S \sin\xi_n + y_{SS})\,\mathrm{tg}\xi_n , \text{ for case BWS;} \tag{26a}$$

$$\overline{C_{SS}F} = r_S \cos\xi_n - x_{SS} + \overline{EF} = r_S \cos\xi_n - x_{SS} + \overline{ET}\,\mathrm{tg}\xi_n = \\ = r_S \cos\xi_n - x_{SS} + (r_S \sin\xi_n - y_{SS})\,\mathrm{tg}\xi_n , \text{ for case CWS.} \tag{26b}$$

Finally, the distance of the SS COG from the tipping line t is calculated according to the equations:

$$l_S = \overline{C_{SS}F}\cos\xi_n = \left[x_{SS} + r_S \cos\xi_n + (r_S \sin\xi_n + y_{SS})\mathrm{tg}\xi_n\right]\cos\xi_n , \text{ for case BWS;} \tag{27a}$$

$$l_S = \overline{C_{SS}F}\cos\xi_n = \left[r_S \cos\xi_n - x_{SS} + (r_S \sin\xi_n - y_{SS})\mathrm{tg}\xi_n\right]\cos\xi_n , \text{ for case CWS.} \tag{27b}$$

The intensity of the overturning moment is calculated according to the expressions:

$$M_T = |Z_{G2}|\, l_T , \text{ for ZG2}\neq 0; \tag{28a}$$

$$M_T = \sum_L k_L L z_L , \text{ L=WL,WWL,N,NN,UL,DD, for ZG2=0,} \tag{28b}$$

where L is the intensity of corresponding longitudinal horizontal force, z_L is the applicate of its point of action, while k_L is the indicator of analyzed influence existence, according to Table 2 (if ''0'' stands in the corresponding field, k_L=0; otherwise, k_L=1).

The ratio between the moment of stability:

$$M_S = G_{SS} l_S , \tag{29}$$

and the overturning moment, expressions (28):

$$\nu = \frac{M_T}{M_S} , \tag{30}$$

presents the factor of safety against overturning, whose value must meet the criterion:

$$\nu \geq \nu_{DIN} , \tag{31}$$

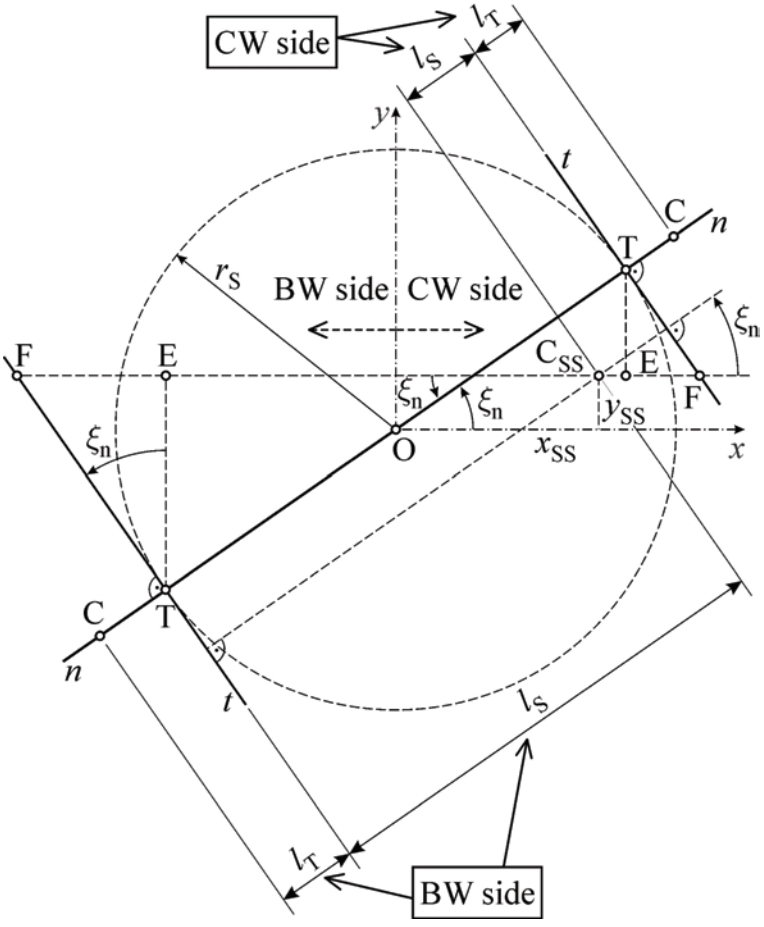


Fig. 5. The scheme for determining the distance of the SS center of gravity (CSS) and penetrating point of the overturning forces principal vector line of action (C) from the line of possible overturning: Oxy-RSB referent plane; n-intersecting line between the overturning moment plane of action and the RSB plane; rS-radius of stability contour; T-the point of intersection between the line n and the stability contour; t-line of possible overturning (tipping line); ξn-angle between intersecting line n and x axis; $\overline{OC}$ =ρC,max; lT-distance between the points C and T; lS-distance between the points CSS and T; E,F-auxiliary points for geometry calculations

Table 3. Positions[a] of points O_1 and O_{BW}

Point	Coordinates		
	x (m)	y (m)	z (m)
O_{BW}[b]	–36.288	–0.76	0.36
O_1[c]	–3.878	0	3.675

[a]Fig. 3.;[b]related to the $O_1x_1y_1z_1$ coordinate system for α_{BWB}=0; [c]related to the Oxyz coordinate system

Table 4. Weights[a] and centers of gravity positions[b] of the SS substructures (forces belonging to G1)

Notation	Weight (kN)	Coordinates		
		x (m)	y (m)	z (m)
G_{BWB}	3992.582	–26.693[c]	–0.35	0.931
G_{M1}	593.770	–4.656[c]	0	17.624
G_{CWB}	2440.483	26.296[d]	0.05	12.526
G_{M2}	778.561	3.773[d]	–0.219	22.688
G_{SP}	1782.605	0.33[d]	0	1.375
G_{CW}	2168.801	34.123[d]	0	10.075

[a]weights of BWBS, CWBS and winch ropes are added to weights of BWB, M1, M2 and CWB; [b]Fig. 3; [c]related to the $O_1x_1y_1z_1$ coordinate system for α_{BWB}=0; [d]related to the Oxyz coordinate system

Table 5. Intensities and positions[a] of points of application of forces belonging to G2

Notation	Intensity (kN)	x (m)	y (m)	z (m)
F_1	376.1	–18.1[b]	0.85	0.8
V_0	196.6	–36.288[b]	–0.76	0
V_1	37.61	–18.1[b]	0.85	0.8
VV_1	597	–35.3[b]	0.9	2.5
W_L/WW_L				
BWB	92.6	–23.433[b]	0	1.034
M1	16.7	–2.91[b]	0	14.329
CWB	90.7	20.116[c]	0	10.725
M2	20.0	3.878[c]	0	20.661
SP	19.7			
N/NN				
BWB	199.629	–26.693[b]	–0.35	0.931
M1	29.689	–4.656[b]	0	17.624
CWB	122.024	26.296[c]	0.05	12.526
M2	38.928	3.773[c]	–0.219	22.688
SP	89.13	0.33[c]	0	1.375
CW	108.401	34.123[c]	0	10.075
F1	18.805	–18.1[b]	0.85	0.8
V0	9.83	–36.288[b]	–0.76	0
V1	1.881	–18.1[b]	0.85	0.8
VV1	29.85	–35.3[b]	0.9	2.5
U_F/UU_F	505.1/658.802	–42.413[b]	–0.12	0.36
U_L/UU_L	505.1/658.802	–36.288[b]	–0.12	–5.765
DD				
BWB	439.184	–26.693[b]	–0.35	0.931
M1	65.315	–4.656[b]	0	17.624
CWB	268.453	26.296[c]	0.05	12.526
M2	85.642	3.773[c]	–0.219	22.688
SP	196.087	0.33[c]	0	1.375
CW	268.453	34.123[c]	0	10.075
F_1	41.371	–18.1[b]	0.85	0.8
V_0	21.626	–36.288[b]	–0.76	0
V_1	4.137	–18.1[b]	0.85	0.8
VV_1	65.67	–35.3[b]	0.9	2.5

[a]Fig. 3; [b]related to the $O_1x_1y_1z_1$ coordinate system for α_{BWB}=0; [c]related to the Oxyz coordinate system

with v_{DIN} being the value of safety factor prescribed by code [4] for the considered load case.

Procedures for the static stability proof 'a priori' and 'a posteriori' are the same, but for conducting an 'a posteriori' procedure it is first of all necessary to determine the 'corrective mass' and its COG on the basis of weighing results, according to the procedure presented in [5].

3. Numerical example

Calculations and comparison of the obtained results were conducted using the in-house developed software STEX. This software enables the determination of all the variables needed for static stability analysis on the whole domain of the BWB hoisting angle ($-19.52° \leq \alpha_{BWB} \leq 14.1°$). Input data (Tables 3-5) are taken from the final stability calculation (''a priori'') provided by the manufacturer of the BWE 1600 (variant V2 in [5]). The positions of acting points of all the forces acting upon SuS1, related to the coordinate system *Oxyz*,

Table 6. Setting values of the winch rope forces: V2

Rope force	S_A (kN)	S_{AA} (kN)	S_Z (kN)	S_{ZZ} (kN)
Setting value	197.4	182.7	342.5	358.3

Table 7. Position of the 'corrective mass' COG related to the Oxyz coordinate system

BWB position	x_C (m)	y_C (m)
L: $\alpha_{BWB}=-12.9°$	−22.39	0.047
H	−22.824	0.311
Hi: $\alpha_{BWB}=14.1°$	−21.199	0.507

Table 8. 'Corrective mass' COG position related to the moving coordinate system $O_1\xi\eta\zeta$

ξ_C (m)	η_C (m)	ζ_C (m)
18.946	0.288	3.29

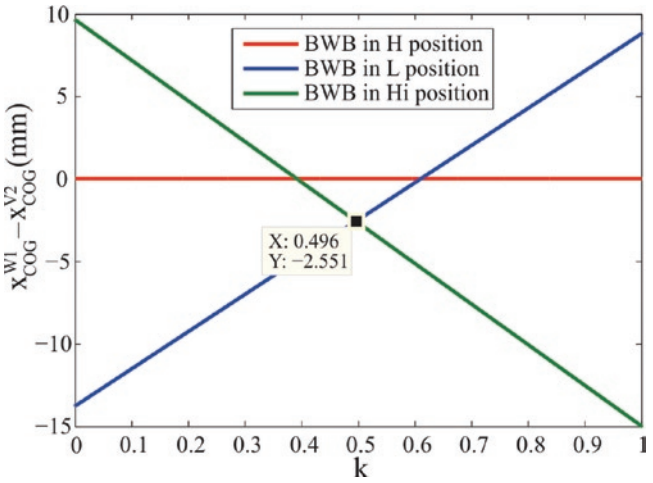


Fig. 9. Deviation of the 'corrective mass' COG abscissa

are determined by the appropriate transformations of coordinates related to the coordinate system $Ox_1y_1z_1$ (Fig. 3).

According to expressions (1)–(4), the maximum and minimum intensities of winch rope forces for LC HZ3.1 ($S_{HZ3.1,max}$=317.1 kN, $S_{HZ3.1,min}$=210.0 kN, Fig. 6), are used as the basis for determining the setting values of the rope forces (Table 6). They determine the intensities of ground reactions when partial leaning of the BW appears, as well as overload forces, equations (6) and (7), Figs. 7 and 8.

Using the procedure and data presented in [5, Subsection 6, expressions (1)–(6)] the 'corrective mass' is determined as follows:

$$\Delta m_S^{V2} = m_{S,W1,A} - m_{S,CW1}^{V2} = 1172.263 - 1154.387 = 17.876 \text{ t},$$

as well as its COG in measuring positions of the BWB:

$$\chi_{C,i} = \frac{G_{S,i}^{W1} x_{COG,i}^{W1} - m_{S,CW1}^{V2} g x_{COG,i}^{V2}}{\Delta m_S^{V2} g}, \ \chi = x, y,$$

where i=L, H, Hi presents the indicator of the BWB measuring posi-

Table 9. BPSS: V2 vs. V5 vs. W1[a]

Variant	Total mass	CW mass	Total mass without CW	x_{COG} (mm)		
				L	H	Hi
	(t)			α_{BWB} (°)		
				−12.9	0	14.1
V2	1154.387	177.017	977.370	−51	−15	581
V5	1172.263	177.017	995.246	−396	−356	252
W1	1172.263	177.017	995.246	−398	−356	249

[a] *Weighing 1 [5]*

Table 10. Distances of penetration points A and B from the slewing axis (BWS)

Load case	$\rho_{A,max}$ (m)		$\rho_{B,max}$ (m)	
	V2	V5	V2	V5
H1.1	2.44	2.714	4.027	4.271
HZ3.1	2.603	2.874	3.717	3.967
HZ3.2	0.564	0.878	1.134	1.442
HZS4.4	4.236	4.475	5.162	5.384
HZS4.8	4.495	4.723	5.504	5.714
HZG5.3	5.092	5.310	5.643	5.851
HZG5.4	3.334	3.604	3.753	4.017

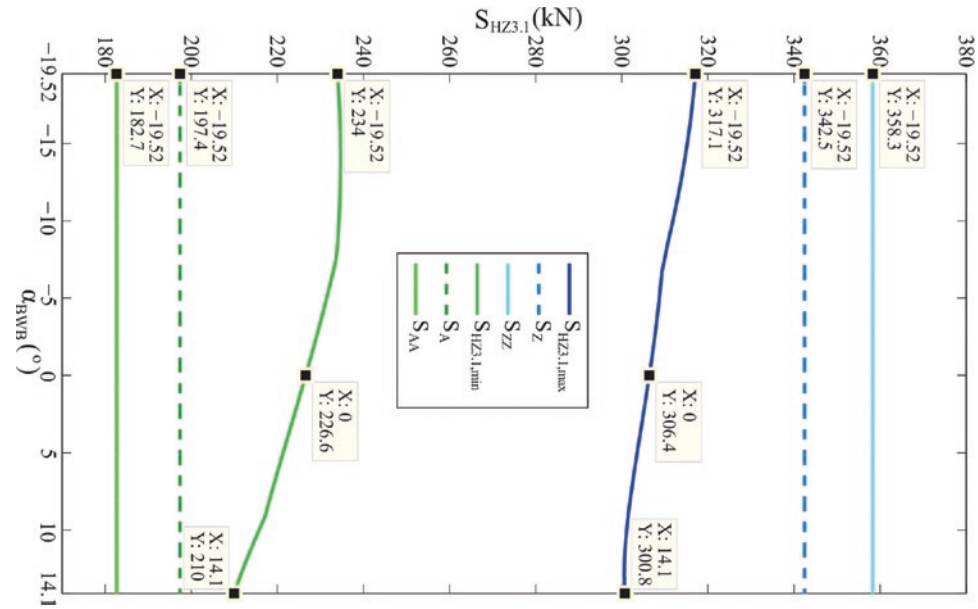


Fig. 6. Rope force in LC HZ3.1 and setting values of the winch rope forces: V2

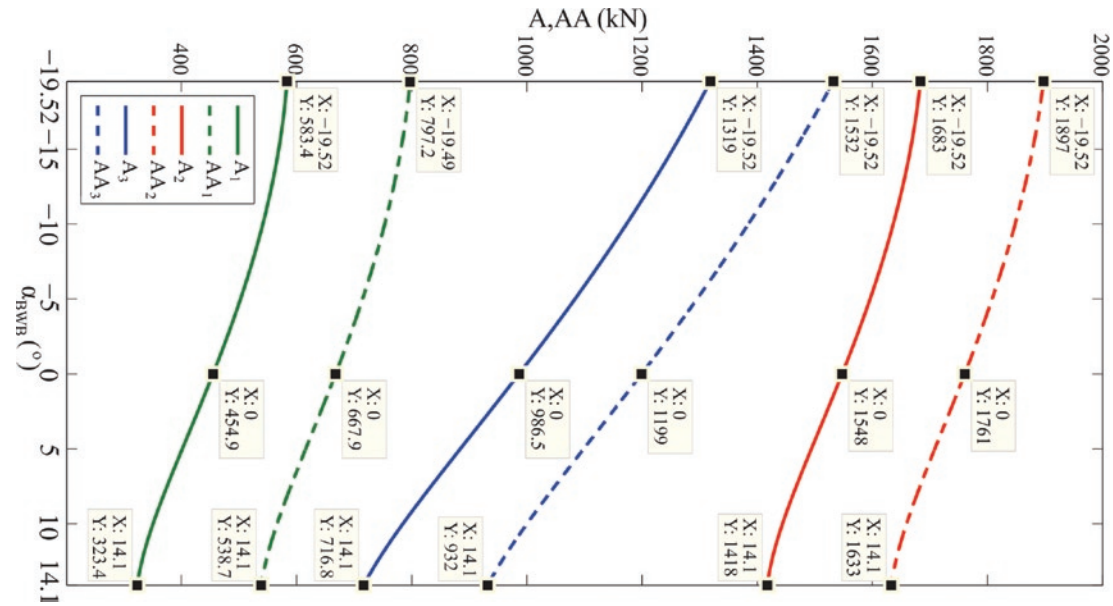


Fig. 7. Ground reactions due to partial leaning of the BW: V2 (for LC 1, LC 2 and LC 3)

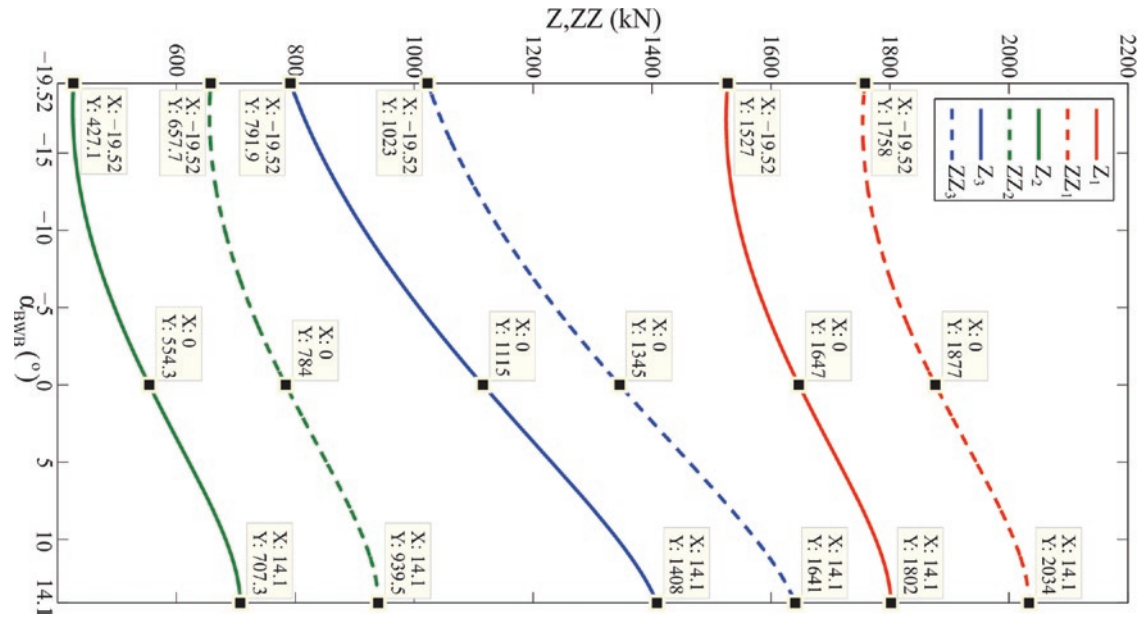


Fig. 8. Overload forces: V2 (for LC 1, LC 2 and LC 3)

Table 11. Distances of penetration points A and B from the slewing axis (CWS)

Load case	$\rho_{A,max}$ (m)		$\rho_{B,max}$ (m)	
	V2	V5	V2	V5
H1.1	2.657	2.317	3.077	2.735
HZ3.1	2.890	2.585	3.245	2.899
HZ3.2	3.124	2.819	3.556	3.206
HZS4.1	3.676	3.379	4.048	3.699
HZS4.7	4.149	3.739	4.739	4.316
HZG5.2	4.974	4.543	5.363	4.923
HZG5.4	3.870	3.531	4.075	3.736

Table 12. Minimum safety factors and corresponding BWB hoisting angles and angles of possible overturning planes (BWS)

Load case	ν_{min}		α_{BWB} (°)		ξ_n (°)	
	V2	V5	V2	V5	V2	V5
H1.1	1.906	1.827	−1.1	−1.34	−0.18	−0.18
HZ3.1	1.806	1.732	0.03	−0.26	−0.17	−0.17
HZ3.2	3.628	3.489	12.9	12.05	0.28	0.28
HZS4.4	1.214	1.165	−1.66	−1.9	−0.55	−0.55
HZS4.8	1.144	1.100	7.25	6.83	0.98	0.97
HZG5.3	1.024	0.985	6.42	5.97	0.99	0.99
HZG5.4	1.463	1.401	5.0	4.6	−0.14	−0.14

Table 13. Minimum safety factors and corresponding BWB hoisting angles and angles of possible overturning planes (CWS)

Load case	ν_{min}		α_{BWB} (°)		ξ_n (°)	
	V2	V5	V2	V5	V2	V5
H1.1	4.056	4.476	14.1	14.1	0	0
HZ3.1	3.174	3.51	14.1	14.1	0	0
HZ3.2	2.606	2.887	14.1	14.1	0	0
HZS4.1	1.834	2.025	14.1	14.1	0	0
HZS4.7	1.363	1.508	−19.52	−19.52	0.05	0.05
HZG5.2	1.066	1.179	−19.52	−19.52	0.04	0.04
HZG5.4	1.66	1.827	14.1	14.1	0	0

Table 14. BPSS. V5 vs. W2 (α_{BWB}=−11.4°, m_{CW}=231.977 t)

Variant	CW mass (t)	Total mass (t)	Coordinates of the COG (mm)	
			x_{COG}	y_{COG}
V5	231.977	1227.223	1130	−116
W2	231.977	1233.772	1087	−130

tion (Table 7).

The abscissa and the applicate of the 'corrective mass' COG related to the moving coordinate system $O_1\xi\eta\zeta$ are determined according to the expressions given in [3, equations (7)–(9)] with the adopted value of the corrective factor k=0.496, while the ordinate is determined as the average of ordinates obtained for the analyzed measuring positions (Table 8).

When transforming V2 to V5 in the paper [5] during the process of determination of the applicate of the 'corrective mass' COG, the value of the corrective factor k=0.5 was adopted and used. In the analysis to follow, a more precise value of the mentioned factor of k=0.496 was used (Fig. 9, Table 9).

The radius of the slew bearing is 5.5 m, so, according to the expression (21), the radius of the stability contour is

$$r_S = 0.95 r_{SB} = 0.95 \times 5.5 = 5.225 \text{ m}.$$

Positions of the penetration points A and B, as well as the values of the factor of safety against overturning, calculated 'a priori' (V2) and 'a posteriori' (V5), are presented in Tables 10−13. Dependence of the factors of safety on the BWB hoisting angle for critical load cases is shown in Figs. 10−13.

Control weighing (W2) was conducted after the correction of the CW mass, with CW of 231.977 t [5, Subsection 5], Table 14.

The weight of the 'corrective mass' causes changes of winch rope force intensity making it necessary to correct the setting values ac-

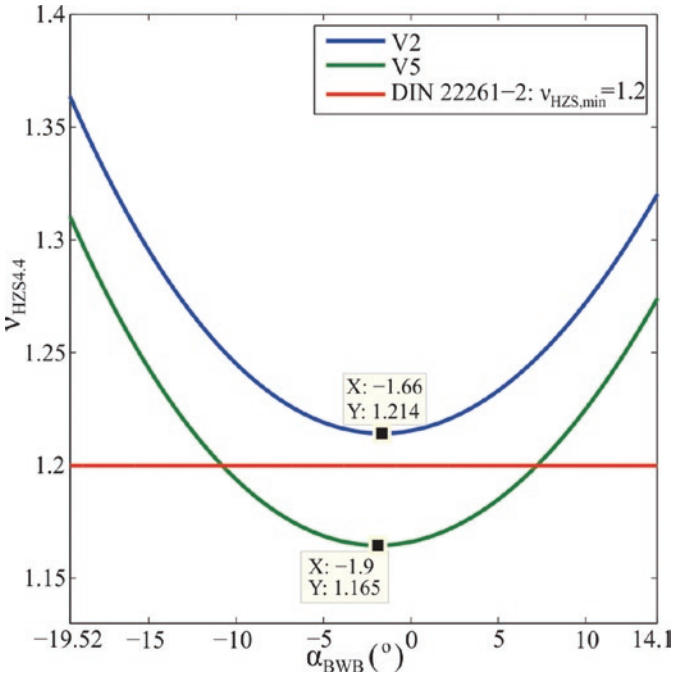


Fig. 10. Safety factor in the LC HZS4.4 (BWS)

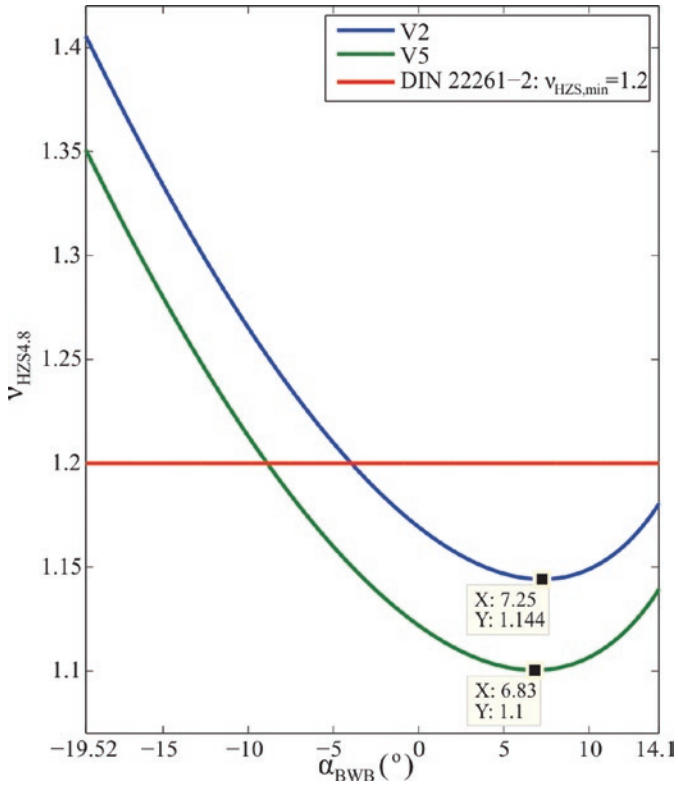


Fig. 11. Safety factor in the LC HZS4.8 (BWS)

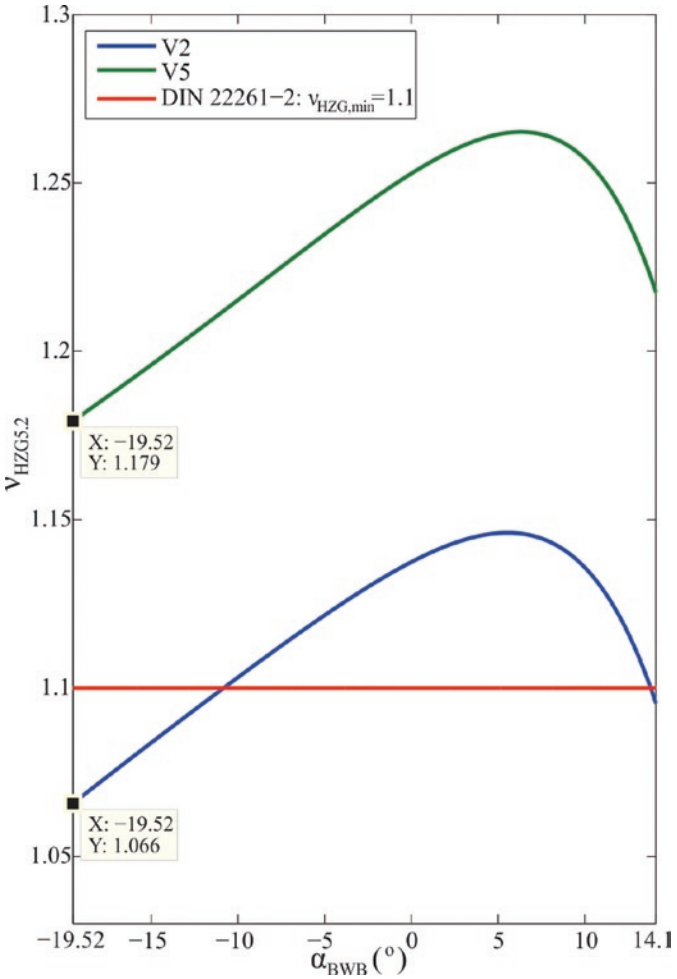


Fig. 12. Safety factor in the LC HZG5.2 (CWS)

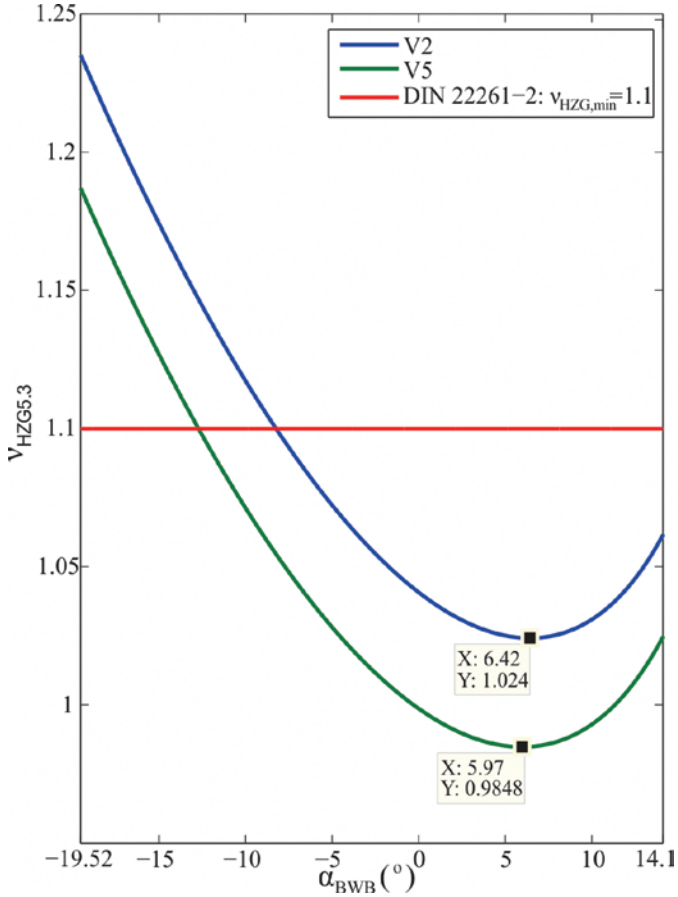


Fig. 13. Safety factor in the LC HZG5.3 (BWS)

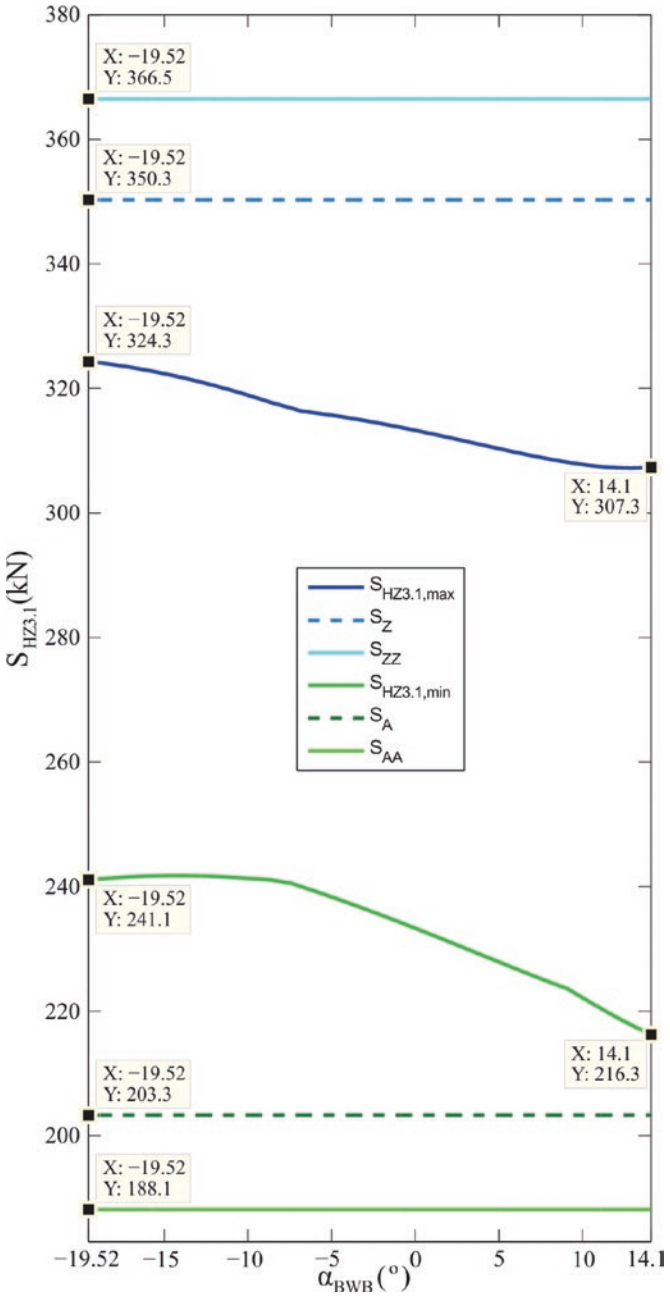


Fig. 14. Setting values of winch rope forces: V5

cording to equations (1)−(4), (Fig. 14, Table 15). The modification of the CW mass in relation to the designed one, caused changes of positions of the points *A* and *B* (Tables 16 and 17) as well as changes of safety factor values (Tables 18 and 19).

4. Discussion

The SS for V2 fulfills the necessary condition of static stability. For all the referent load cases, the distance of penetration point *A* from the slewing axis (Tables 10 and 11) fulfills the condition defined by expression (22). The maximum distance value of $\rho_{A,max}$=5.092 m which is obtained for the LC HZG5.3 (BWS) is lower than the value of the stability contour radius defined by the expression (21), r_S=5.225 m. However, in LCs HZS4.8, HZG5.2 and HZG5.3 the distances of point *B* from the slewing axis are greater than the radius of the stability contour, which means that in the described load cases the condition defined by the expression (23) is not fulfilled. Consequently, in the above-mentioned LCs, safety factors are lower than those prescribed by standard [8] (Tables 12 and 13, Figs. 10−13).

The 'corrective mass' and its COG position were determined on the basis of measurement 1 conducted after the first erection (Table 9) using the procedure presented in [5, Subsection 2]. Variant V2 was transformed into variant V5 by implementing the influence of the corrective mass. By adopting the corrective factor value of k=0.496, the absolute values of COG abscissa deviations, as opposed to those determined by measurement 1, become lower than 3 mm (Fig. 9, Table 9). Variant V5, formed in that manner, presents the actual image of the SS, as opposed to the V2, which is the designed (desired) image of the SS.

The excessive mass ('corrective mass') on the side of the BW causes deviation of the penetration points *A* and *B* distances from the slewing axis (Tables 20 and 21). An increase of referent distances (Table 20) worsens the stability conditions on the BW side, leading to lower values of safety factors (Table 12) as well as widening of the BWB hoisting angle range within which the values of safety factors are lower than allowed (Figs. 10 and 13). Apart from the already mentioned LCs HZ4.8 and HZG5.3, the value of safety factor for V5 is also lower than the minimum value prescribed by standard [8] in the LC HZS4.4 (Fig. 10). On the other hand, in the CWS case, the decrease of referent distances improves the stability conditions, thus leading to greater safety factor values (Table 13). As opposed to V2, the SS in V5 also meets the static stability criterion in LC HZG5.2 (Fig. 12).

$$\Delta x_{COG} = x_{COG}^{W1} - x_{COG}^{V2}$$

Table 15. Setting values of winch rope forces: V5

Rope force	S_A (kN)	S_{AA} (kN)	S_Z (kN)	S_{ZZ} (kN)
Setting value	203.3	188.1	350.3	366.5

Table 16. Distances of penetration points A and B from the slewing axis: V5, m_{CW}=231.977 t (BWS)

Load case	$\rho_{A,max}$ (m)	$\rho_{B,max}$ (m)
H1.1	2.416	3.975
HZ3.1	2.576	3.669
HZ3.2	0.58	1.139
HZS4.4	4.177	5.087
HZS4.8	4.715	5.746
HZG5.3	5.305	5.866
HZG5.4	3.309	3.722

Table 17. Distances of penetration points A and B from the slewing axis: V5, m_{CW}=231.977 t (CWS)

Load case	$\rho_{O,max}$ (m)	$\rho_{O,max}^{G2,fac}$ (m)
H1.1	2.606	3.023
HZ3.1	2.876	3.184
HZ3.2	3.107	3.488
HZS4.1	3.667	3.982
HZS4.7	3.715	4.218
HZG5.2	4.529	4.872
HZG5.4	3.819	4.023

Table 18. Minimum safety factors and corresponding BWB hoisting angles and angles of possible overturning planes: V5, m_{CW}=231.977 t (BWS)

Load case	ν_{min}	α_{BWB} (°)	ξ_n (°)
H1.1	1.931	−1.12	−0.18
HZ3.1	1.83	0	−0.17
HZ3.2	3.668	12.78	0.28
HZS4.4	1.232	−1.7	−0.55
HZS4.8	1.098	6.75	0.98
HZG5.3	0.986	5.93	0.99
HZG5.4	1.476	5.04	−0.14

Table 19. Minimum safety factors and corresponding BWB hoisting angles and angles of possible overturning planes: V5, m_{CW}=231.977 t (CWS)

Load Case	ν_{min}	α_{BWB} (°)	ξ_n (°)
H1.1	4.142	14.1	0
HZ3.1	3.252	14.1	0
HZ3.2	2.677	14.1	0
HZS4.1	1.875	14.1	0
HZS4.7	1.591	−19.52	2.87
HZG5.2	1.202	−19.52	2.46
HZG5.4	1.688	14.1	0

Table 20. Differences of the penetration points A and B distances from the slewing axis (BWS)

Load case	$\rho_{A,max}^{V5} - \rho_{A,max}^{V2}$ (m)	$\rho_{B,max}^{V5} - \rho_{B,max}^{V2}$ (m)
H1.1	0.27	0.244
HZ3.1	0.271	0.25
HZ3.2	0.314	0.308
HZS4.4	0.239	0.222
HZS4.8	0.228	0.21
HZG5.3	0.218	0.208
HZG5.4	0.27	0.264

Mass distribution of the actual SS is less favorable than for the designed one, which causes shifting of the SS center of gravity towards the BW (Tables 9 and 22).

The excavator designer solved the problem of unfavorable COG shifting towards the BW by enlarging the CW mass. Based on the results of measurement 1, the mid value of the excessive moment was calculated in relation to the designed values (V2) and then divided by the distance of the CW COG from the slewing axis. That way, the weight which has to be added to the designed weight of the CW was calculated.This is the reason why the mass of the CW during the measurement 2 was 231.977 t, as opposed to the design one (V2) which was 221 t (Table 14). The difference between the SS mass determined by measurement 2 and the one obtained by V5 (6.549 t, Table 14) is predominantly the consequence of the appearance of foreign bodies and a little bit of snow [5]. Having in mind the conditions in which measurement 2 was conducted, the calculated coordinates of COG for V5 are in good correlation to those obtained by measurement 2 (Table 14).

Greater weight on the BW side, as opposed to the designed one, led to the increase of winch rope force intensity (Figs. 6 and 14) and

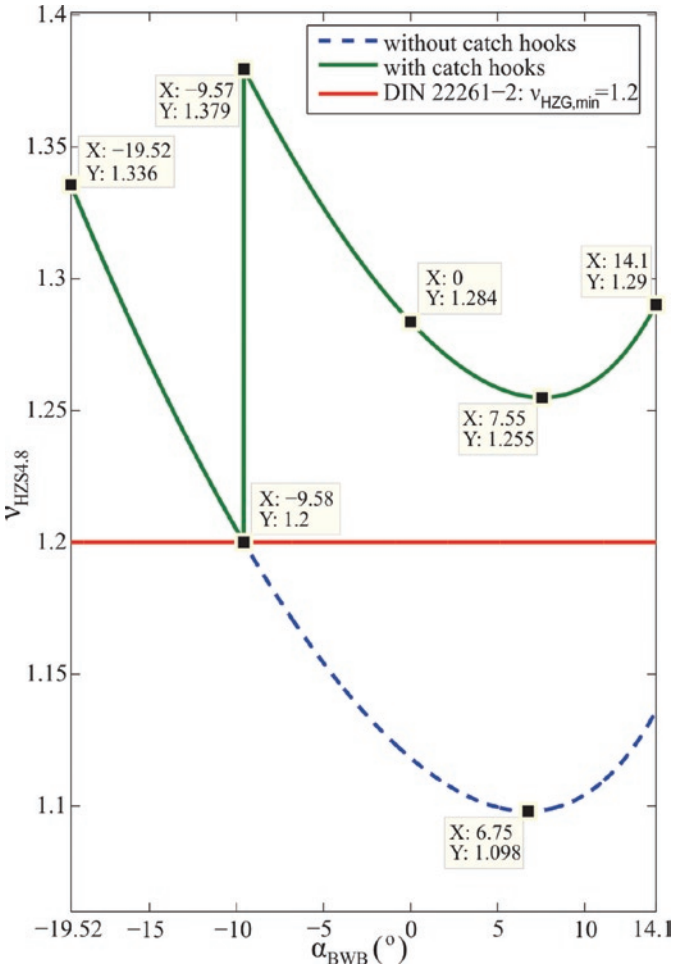


Fig. 15. Safety factor in LC HZS4.8 (BWS): V5 (m_{CW}=231.977 t)

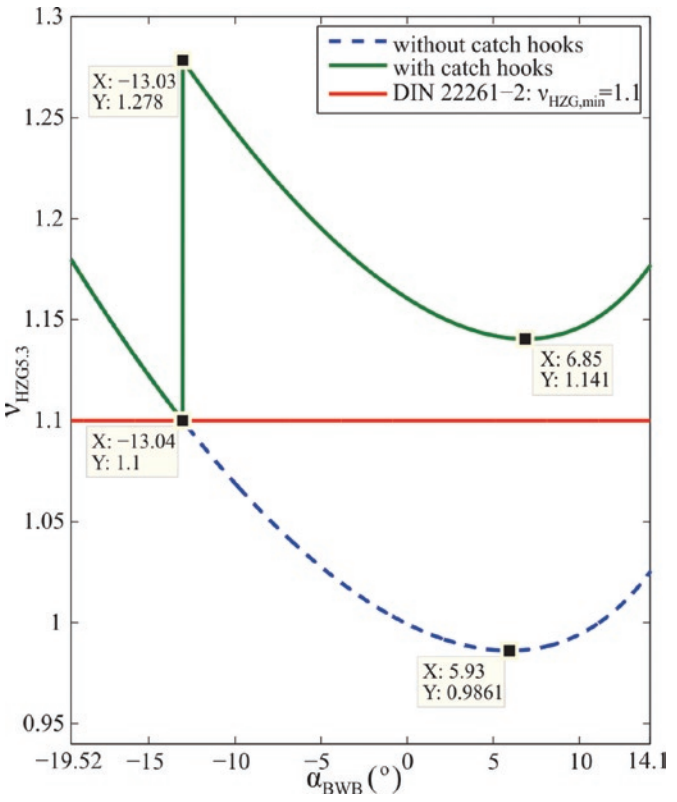


Fig. 16. Safety factor in LC HZG5.3 (BWS): V5 (m_{CW}=231.977 t)

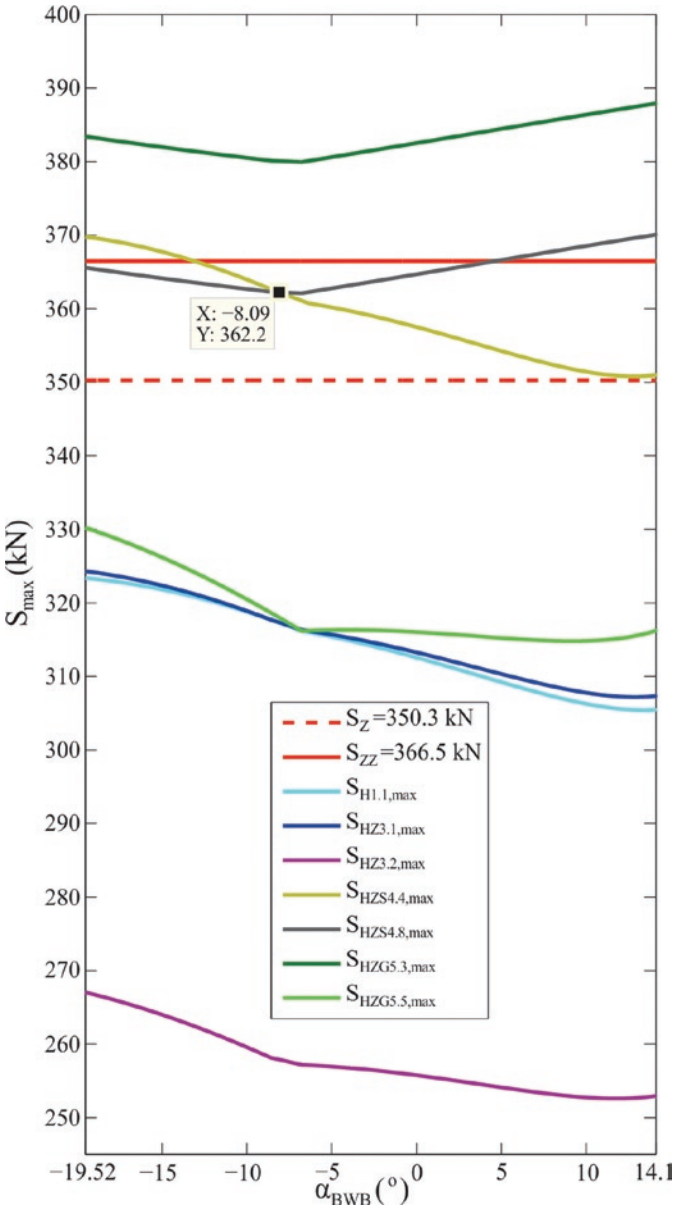


Fig. 17. Maximum winch rope forces: V5

accordingly, to the increase of setting values (Tables 6 and 15). The setting values of minimum forces are higher for ≈3.0%, while setting values of maximum winch rope forces are ≈2.3% higher.

With corrected CW mass (231.977 t) the SS fulfills the necessary condition of static stability, expression (21), as well as the sufficient one, expression (23), on the CW side (Table 17). This is confirmed by the obtained safety factor values for all the analyzed load cases (Table 19) which are higher than the minimum ones prescribed by the code [8].

When it comes to the static stability on the BW side, it can be concluded that, for all the load cases except LC HZG5.3 (Table 16) the necessary static stability condition is fulfilled. Apart from the above-mentioned load case, the sufficient static stability condition is also not fulfilled in LC HZS4.8 (Tables 16 and 18). The eventual loss of static stability in the mentioned critical load cases is prevented by activation of catch hooks (Figs. 15 and 16).

It is important to note that the SS in LC HZS4.4 fulfills the prescribed static stability condition [8], although the intensity of the maximum winch rope force is higher than that obtained for LC HZS4.8 on the BWB hoisting angle interval of $-19.52° \leq \alpha_{BWB} \leq -8.09°$ (Fig. 17).

Table 21. Differences of the penetration points A and B distances from the slewing axis (CWS)

Load case	$\rho_{A,max}^{V5} - \rho_{A,max}^{V2}$ (m)	$\rho_{B,max}^{V5} - \rho_{B,max}^{V2}$ (m)
H1.1	−0.34	−0.342
HZ3.1	−0.305	−0.346
HZ3.2	−0.305	−0.35
HZS4.1	−0.297	−0.349
HZS4.7	−0.41	−0.423
HZG5.2	−0.431	−0.44
HZG5.4	−0.339	−0.339

Table 22. Difference between the 'measured' and the projected abscissas of the SS center of gravity

Difference between GOG abscissas	BWB measuring position		
	L: α_{BWB}=−12.9°	H	Hi: α_{BWB}=14.1°
(mm)	−347	−341	−332

Table 23. Minimum safety factors and the belonging BWB hoisting angles: V5 (y_i=0), m_{CW}=231.977 t

Load case	BWS		CWS	
	ν_{min}	α_{BWB} (°)	ν_{min}	α_{BWB} (°)
H1b	1.931	−1.12	4.142	14.1
HZ2	1.83	0	3.252	14.1
HZ3	3.668	12.78	2.677	14.1
HZS4			1.875	14.1
HZS5	1.231	−1.7		
HZS8			1.59	−19.52
HZS9	1.099	6.75		
HZG13			1.202	−19.52
HZG14	0.986	5.93		
HZG15	1.476	5.04	1.688	14.1

This is explained by the influence of the SS geometric configuration, primarily by the mutual position of axis O_1y_1 and tipping line t (Figs. 3 and 5), which is decisive for the calculation of intensities of winch rope forces and moments of overturning and stability, and also by the positions of loads in the discussed load cases.

Having in mind the fact that the angles ξ_n defining the plane of overturning are relatively small (Tables 12, 13, 18 and 19), in the analysis that follows, the presented results, (Table 23), were obtained by introducing the presumption that the eccentricities of all forces acting upon the SS, in the relation to the plane *Oxz* (Tables 4 and 5) are equal to zero (i.e. y_i=0).

Based on the comparative analysis of results (Tables 18, 19 and 23), it can be concluded that, for the adopted level of accuracy, the calculation model with y_i=0 gives equal values of safety factors and values of corresponding BWB hoisting angles, as the one including the influence of loads eccentricities in relation to the *Oxz* plane.

In engineering practice, it is common to conduct the static stability calculation only for characteristic BWB positions: L (α_{BWB}=−19.52°), H (α_{BWB}=0) and Hi (α_{BWB}=14.1°). Calculation results (Tables 13, 19 and 23) show that this kind of approach might be used when dealing with the stability on the side of the CW. On the other hand, when the stability on the BW side is in question, it can be observed that there are some load cases within the minimal safety factor values obtained for BWB positions which are neither close to the H nor to the Hi position (Tables 12, 18 and 23). For example, in LC HZS4.8, the minimal value of safety factor of ν_{min}=1.255 (with catch hooks) is obtained for the BWB hoisting angle of α_{BWB}=7.55° (Fig. 15). The mentioned value is ≈2.3% and ≈2.8% lower than those obtained for positions H and Hi, respectively.

5. Proposal of the superstructure static stability proof procedure 'a posteriori'

According to the literature which was available to the authors, the answer to the question „what to do after the superstructure weighing" is given only in [9, page 233]: „If the weighing results differ by more than a certain amount, in general 5% of the theoretical values calculated for stability, the calculation must be checked and the weighing procedure repeated. The ballast must then be adjusted according to the weighing results so that the position of the COG in the plane of the jacking points corresponds to the desired theoretical values.". Therefore, the answer is of a general character. Having in mind the extreme importance of the static stability problem, considering the presented procedures and obtained investigation results, the authors suggest the implementation of the static stability calculation procedure 'a posteriori'. Its fundamental phases are:

1. Harmonization of the calculation model with all the changes made during the development of the project and the realization

of the first erection procedure, i.e. forming of the calculation model 'a priori';
2. Conducting the first weighing immediately after the completion of the first erection procedure;
3. Calculation of weight and COG position of 'a priori' model in measuring positions of the BWB;
4. Calculation of the 'corrective mass' and its position based on the results obtained by the first measurement (weighing);
5. Forming of the 'a posteriori' model, taking into account the influence of the 'corrective mass';
6. Validation of the calculation 'a posteriori' model using the results of the first measurement;
7. Determination of the needed correction of the CW mass and setting values of winch rope forces based on the static stability calculation conducted on the 'a posteriori' model;
8. Second (control) measurement with the corrected mass of the CW;
9. Calculation of weight and COG position of the 'a posteriori' model for positions of the BWB during the second measurement;
10. Final validation of the 'a posteriori' model based on the results of the second measurement;
11. Final calculation of minimal values of factor of safety against overturning on the complete domain of the BWB hoisting angle, using the 'a posteriori' model.

6. Conclusion

The complexity of the slewing superstructure balancing problem, hence the determination of the CW mass, is the consequence of: (a) the changeability of the geometric configuration; (b) the complexity of working conditions and (c) multiple limitations of the possible set of solutions. Namely, the mass of the CW has to be determined in such a manner to fulfill the prescribed criteria of static stability in all working conditions and for all geometric configurations. Furthermore, in load cases matching the normal excavator operating conditions (LCs H1.1b), the penetration point of the total superstructure loading principal vector line, through the referent plane of the slew bearing, must not be more than one quarter of its diameter from the slewing axis. That would guarantee a reliable and long-term operation of the slew bearing. Finally, from the upper side, the mass of the CW is limited by the CWB structure carrying capacity.

Keeping in mind the fact that the existing referent literature does not specify the procedure of static stability proof, this paper is the first to present it in detail. The transformation of the calculation model 'a priori' (designed image of the superstructure) to the calculation model 'a posteriori' (actual image of the superstructure) was conducted on the basis of weighing results. Calculation of all the parameters that determine the BWE superstructure static stability, for both calculation models, was conducted using the original software, developed on the basis of the above-mentioned procedure. Based on the presented calculation results, the following conclusions are derived:

- with the designed mass of the CW, the 'a priori' model fulfills the necessary condition of the static stability in all analyzed LCs, which was not the case with the 'a posteriori' model;
- with the designed mass of the CW, both models do not meet the sufficient criterion of the static stability in three LCs, two of which are the same;
- with the corrected mass of the CW, the 'a posteriori' model does not fulfill the necessary condition of the static stability in one LC, and sufficient criterion in two LCs;
- with the corrected mass of the CW, the 'a posteriori' model, which is of a spatial nature, may be reduced to the planar model of high results accuracy.

A particular contribution, achieved on the basis of the presented investigation results and perennial experience, is represented in the classification of the SS models into two fundamental groups which were named the 'a priori' and the 'a posteriori' models, as well as the fact that the basic stages of the 'a posteriori' model forming are presented in a paper. The 'a posteriori' models enable a reliable calculation of the SS static stability and may be used not only for static stability proof of the machine as a whole, but also for loads analysis of substructures and elements of BWEs and related surface mining machines – spreaders. Moreover, previously-mentioned models are of extreme importance for a successful and reliable exploitation of the machine since they present the basis for:

- adjustment and control of limiting winch rope forces values;
- periodic control of mass and center of gravity position;
- eventual reconstruction, which would be conducted in order to realize better customization of the machine versus operating conditions;

which is of extreme importance having in mind the fact that bucket wheel excavators and spreaders are machines intended (designed) for perennial exploitation.

Acknowledgement

This work is a contribution to the Ministry of Education, Science and Technological Development of Serbia funded project TR 35006.

References

1. Augustynowicz J, Dudek K, Figiel A, Nowak J, Kluczkiewicz W. Doświadczalne wyznaczanie położenia środka ciężkości obrotowego nadwozia koparek kołowych. Górnictwo Odkrywkowe 2011; 52(3-4): 92-95.
2. Bošnjak S, Zrnić N. Dynamics, failures, redesigning and environmentally friendly technologies in surface mining systems. Archives of Civil and Mechanical Engineering 2012; 12(3): 348-359, https://doi.org/10.1016/j.acme.2012.06.009.
3. Bošnjak S, Petković Z, Dunjić M, Gnjatović N, Đorđević M. Redesign of the vital subsystems as a way of extending the bucket wheel excavators life. Technics Technologies Education Management 2012; 7(4): 1620-1629.
4. Bošnjak S, Oguamanam D, Zrnić N. The influence of constructive parameters on response of bucket wheel excavator superstructure in the out-of-resonance region. Archives of Civil and Mechanical Engineering 2015; 15(4): 977-985, https://doi.org/10.1016/j.acme.2015.03.009.
5. Bošnjak S, Gnjatović N, Savićević S, Pantelić M, Milenović I. Basic parameters of the static stability, loads and strength of the vital parts of a bucket wheel excavator's slewing superstructure. Journal of Zhejiang University - SCIENCE A (Applied Physics & Engineering) 2016; 17(5): 353-365, https://doi.org/10.1631/jzus.A1500037.
6. Brkić A Đ, Maneski T, Ignjatović D, Jovančić P, Spasojević Brkić V K. Diagnostics of bucket wheel excavator discharge boom dynamic performance and its reconstruction. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16 (2): 188-197.
7. Bugaric U, Tanasijevic M, Polovina D, Ignjatovic D, Jovancic P. Lost production costs of the overburden excavation system caused by rubber belt failure. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14 (4): 333-341.
8. DIN 22261-2. Bagger, Absetzer und Zusatsgeräte in Braunkohlentagebauen. Teil 2: Berechnungsgrundlagen. Deutsches Institut für Normung;

2006.
9. Durst W, Vogt W. Bucket Wheel Excavator, Clausthal-Zellerfeld: Trans Tech Publications, 1988.
10. Kowalczyk M, Czmochowski J, Rusiński E. Construction of diagnostic models of the states of developing fault for working parts of the multi-bucket excavator. Eksploatacja i Niezawodnosc- Maintenance and Reliability 2009; 42(2): 17-24
11. Maslak P, Przybylek G, Smolnicki T. Comparison of selected methods for the determination of the center of gravity in surface mining machines. Materials Today: Proceedings 2017: 4(5, Part 1): 5877- 5882, https://doi.org/10.1016/j.matpr.2017.06.062.
12. Nan N, Kovacs I, Popescu F. Balance control by weighting and tensiometric measurements of bucket wheel excavators. WSEAS Transactions on Systems and Control 2008; 3(11): 927-938.
13. Pietrusiak D. Evaluation of large-scale load-carrying structures of machines with the application of the dynamic effects factor. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (4): 542-551, https://doi.org/10.17531/ein.2017.4.7.
14. Pietrusiak D, Smolnicki T, Stańco M. The influence of superstructure vibrations on operational loads in the undercarriage of bulk material handling machine. Archives of Civil and Mechanical Engineering 2017; 17(4): 855-862, https://doi.org/10.1016/j.acme.2017.03.001.
15. Rusiński E, Czmochowski J, Moczko P, Pietrusiak D. Surface Mining Machines - Problems of Maintenance and Modernization, Cham: Springer International Publishing AG, 2017, https://doi.org/10.1007/978-3-319-47792-3.
16. Smolnicki T, Stańco M, Pietrusiak, D. Distribution of loads in the large size bearing - problems of identification. Tehnički Vjesnik - Technical Gazette 2013: 20(5): 831-836.
17. Smolnicki T, Pękalski G, Jakubik J, Harnatkiewicz P. Investigation into wear mechanisms of the bearing raceway used in bucket wheel excavators. Archives of Civil and Mechanical Engineering 2017: 17(1):1-8, https://doi.org/10.1016/j.acme.2016.07.008.

Srđan BOŠNJAK
University of Belgrade
Faculty of Mechanical Engineering
Kraljice Marije 16, 11120 Belgrade 35, Serbia

Nebojša B. GNJATOVIĆ
University of Belgrade
Faculty of Mechanical Engineering
Kraljice Marije 16, 11120 Belgrade 35, Serbia

Ivan MILENOVIĆ
University of Belgrade
Faculty of Mechanical Engineering
Kraljice Marije 16, 11120 Belgrade 35, Serbia

E-mails: sbosnjak@mas.bg.ac.rs, ngnjatovic@mas.bg.ac.rs, imilenovic@mas.bg.ac.rs

Article citation info:
ZHU L, ZHANG Y, ZHANG R, ZHANG P. Time-dependent reliability of spur gear system based on gradually wear process. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 207–218, http://dx.doi.org/10.17531/ein.2018.2.05.

Lisha ZHU
Yimin ZHANG
Rui ZHANG
Peiming ZHANG

TIME-DEPENDENT RELIABILITY OF SPUR GEAR SYSTEM BASED ON GRADUALLY WEAR PROCESS

ZALEŻNA OD CZASU NIEZAWODNOŚĆ UKŁADU PRZEKŁADNI ZĘBATEJ JAKO FUNKCJA PROCESU STOPNIOWEGO ZUŻYCIA

To study dynamic evolution law of mechanical reliability caused by wear, gear transmission system is taken as a research object. Considering the effect of clearance caused by wear on gear teeth load in double meshing area, the formula of dynamic distribution load which is undertaken by two adjacent teeth is deduced. And the distributed pressure and meshing speed, which should be taken into account while calculating gear wear, are obtained based on the Winkler surface model and principle of tooth mesh. Based on the Archard's wear model, numerical simulation model for wear in spur gear is deduced, and the wear depth of each meshing points on teeth outline with different wear cycles are obtained. The calculation wear model is replaced with a surrogate model with Neural Network and Kriging method to overcome time-consuming defect. Random process model is integrated with the surrogate model, and dynamic reliability for nonlinear stochastic structure with unknown distribution characteristic is obtained with Neural Network-based Edgeworth series technique and four moment methods, which is compared with Kriging-based Monte Carlo simulation method. The computational efficiency and accuracy are also demonstrated.

***Keywords**: reliability; gear; time-dependent; wear; gradually.*

W artykule badano prawo dynamicznej ewolucji niezawodności mechanicznej powodowanej zużyciem na przykładzie układu przekładni zębatej. Na podstawie rozważań nad wpływem luzu powstałego na skutek zużycia na obciążenie zębów przekładni w obszarze podwójnych zazębień, wyprowadzono wzór na dynamiczny rozkład obciążeń przyjmowanych przez pary sąsiadujących zębów. Rozłożone naciski i prędkość zazębiania, które należy uwzględnić przy obliczaniu zużycia przekładni, otrzymano na podstawie modelu powierzchniowego Winklera oraz zasady zazębienia. W oparciu o model zużycia Archarda, wyprowadzono numeryczny model symulacyjny zużycia w przekładni zębatej oraz obliczono głębokość zużycia każdego z punktów zazębienia na zarysie zębów przy różnych cyklach zużycia. Aby uniknąć problemu czasochłonności, obliczeniowy model zużycia zastąpiono modelami zastępczymi bazującymi na sieci neuronowej i metodzie krigingu. Model procesu losowego zintegrowano z modelem zastępczym, a dynamiczną niezawodność dla nieliniowej struktury stochastycznej o nieznanej charakterystyce rozkładu uzyskano za pomocą techniki serii Edgeworth opartej na sieci neuronowej oraz metody czterech momentów, którą porównano z metodą symulacji Monte Carlo opartą na krigingu. Wykazano także wydajność obliczeniową i dokładność omawianej metody.

***Słowa kluczowe**: niezawodność; koło zębate; zależność od czasu; zużycie; stopniowy.*

1. Introduction

Friction and wear are unavoidable for transmitting power in gear systems. Severe wear can cause mechanical component damage. The dynamic characteristic of gear system is affected not only by friction and wear but also by variation of geometrical shape and dimension parameter caused by mild wear, which lead to more serious wear. Therefore, the coupled relation between wear and dynamic load should be taken into consideration to study gear wear.

Gear parameters might not be accurate after machining in practice, which is one of reasons why product failure still occurs. That is to say, parameters have uncertainty and randomness. Besides, wear depth accumulates over time gradually, so gear wear is a random process, which has random statistical regularity. As a result, more practical surface wear model and dynamic reliability model with gradually parameters can be established from the definition of reliability, the reliability of mechanical component can be observed all over its life cycle, and the tolerance of designed parameters can be determined to reduce failure probability.

The wear model should be established based on failure mechanism to study gear wear. In recent 30 years, wear and failure mechanism of gears have been studied based on lots of laboratory tests, Zurowski[28] et al. employed TT-3 tester to research wear resistance of C45（norm）/145Cr6 and C45(600)/145Cr6 matchings and indicated that wear resistance of the two matchings had a significant dependence on friction area temperature and material hardness. However, research on establishing models and calculation methods for gear wear have little been touched on. In general, these models break down into three categories:

The first model was established on undetermined coefficient or regression analysis method according to test data. The conclusion was convincing but only obtained laws for intermediate variables. For example, Põdra[15] studied wear had a linear correlation with normal load and had incomplete correlation with sliding velocity.

The second was developed to discuss the influence of parameters from energy loss aspect. For example, Onishchenko[14] studied the effect of machine operation, corresponding tribological theories, the eccentricity of pitch circle and the instant temperature in the contact

on wear. The model stated that wear is simply proportional to specific power, which needed further study.

The third was the Archard wear model[3] which had been widely adopted. The formula had evolved from severe sliding of intermeshing tooth. Andersson[2] obtained analytical formula of sliding distance which varied with different mesh position. Flodin and Andersson[9] developed a numerical model for wear prediction of spur gear and the contact were modelled by Winkler's elastic foundation model. Park[16] combined Archard's wear model with a finite-element based hypoid gear contact model for simulation of surface wear of hypoid gear pairs. Most of the above adopted equal or linear distributed load[2,4,6], without coupled relations between tooth load and wear. Flodin[8,9] studied the effect of spur gears on contact condition, whereas, the sequence of contact, clearance size between two adjacent teeth and time-varying stiffness were not considered.

The parameters were deterministic in the wear models above, randomness of excited load and gradual wear were also not considered. As a matter of fact, wear-based reliability of mechanical product is varying gradually with time[22]. There are many methods to build time-dependent reliability model, such as Markov theory[19], dynamic fault-tree model[7,13], stochastic MCS model[5], Go-flow model[11], stochastic petri nets[20]. As mechanical structure is so complex that the methods above are hardly applied to dynamic reliability analysis for mechanical systems and components. One is because the relationship between response and parameters are nonlinear and there is no analytic expression for the state limited function in general. The other is lack of statistical data and distribution law for parameters.

Reliability analysis is based on probability analysis and mathematical statistics, so it needs enough samples to determine probability characteristic of variables and response. Therefore, it would be time-consuming to obtain enough samples for complex structure. To establish reliability model for complex structure system, the accurate numerical model is replaced with one of surrogate models based on response surface, neutral network, radial basis function network, support vector machine and kriging method, and reliability are obtained by surrogate-model-based monte carlo simulation method. Gomes and Awruch[10] presented response surface and artificial neural network techniques to solve complex and more elaborated problems and carried out comparison using FORM, direct MCS and MCS with adaptive importance sampling. Tan[18] proposed radial basis function networks and support vector machines for reliability analysis. Zhu and Du[27] applied the kriging-based MCS method which was suitable for highly nonlinear limit-state functions to reduce the computational cost. Because the nonlinear response with complex structure can be rapidly predicted in any design point using a kriging interpolator and the iteration process has higher stabilization and faster convergence speed than other methods, the Kriging model is employed in the present paper.

To solve the problem of unknown distribution, on the basis of the surrogate model, Edgeworth series and Four moment techniques are suitable for reliability analysis. The method combines analysis method and moment method and Zhang[26] developed a statistical fourth moment method to examine reliability of the rotor-stator systems with rubbing. All the models and methods above have not been used in analysing wear-based reliability of gear system.

In the present paper, a numerical wear model for a pair of spur gears is established considering coupled relation of wear and dynamic load. A time-dependent reliability model with gradual parameters is also build under the condition that original parameters are stochastic and wear is gradually changing. The time-varying reliability curves for spur gears are obtained during the life cycle.

2. Load distribution theory

As shown in Figure 1, in one meshing cycle, single teeth-meshing area and double teeth-meshing area appear alternately. In single teeth-meshing area, the transmission load is carried only by one pair of mesh gear. In double teeth-meshing area, the transmission load is carried by two pairs of mesh gear. If only considering elastic deformation without wear, the two pairs of mesh gear can be regarded as parallel springs. The distribution load is calculated according to time-varying stiffness of every pair. However, clearance is formed among teeth pair caused by wear in gears, which are run in low speed and heavy load, thus meshing point will deviate from its normal involute location. So, it is necessary to consider time-varying meshing stiffness, the sizes of the two clearances and the contact sequence of the two teeth pairs.

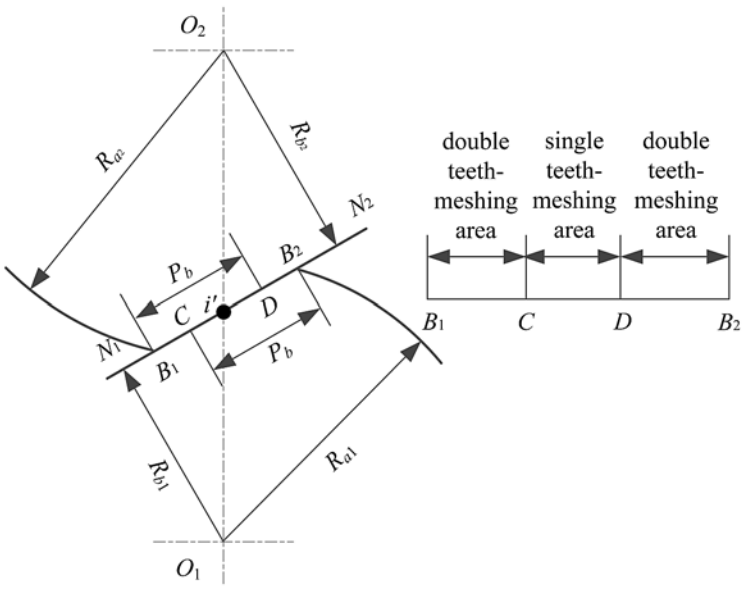


Fig. 1. Schematic diagram of a meshing gear

Take the contact ratio $1<\varepsilon<2$, for instance, as shown in Figure 1. The total transmission load equals the sum of load carried by two adjacent teeth pairs, that is:

$$W = W_1 + W_2 \tag{1}$$

Where W is total normal load per unit width, W_1 is normal load shared by the first gear pair, W_2 is normal load shared by the second gear pair.

As shown in Figure 2, In case of $e_{t1}<e_{t2}$, the first gear pair contacts firstly, e_{t1} and e_{t2} are composite error of gear pair 1 and 2, where the direction along gear surface dent is positive. δ_1 and δ_2 are elastic deformation of the two gear pairs. $k_1(t)$ and $k_2(t)$ are meshing stiffness per unit width.

According to the principle that the load is proportional to the deformation, W_1 and W_2 can be established as:

$$W_1 = k_1(t) \times (x - e_{t1}) \tag{2}$$

$$W_2 = k_2(t) \times (x - e_{t2}) \tag{3}$$

Load distribution coefficient β can be obtained by equations (1) to (3), that is:

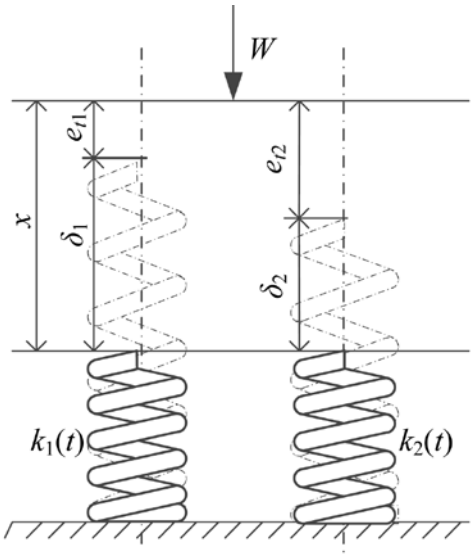


Fig. 2. Force diagram of meshing gears (condition 1)

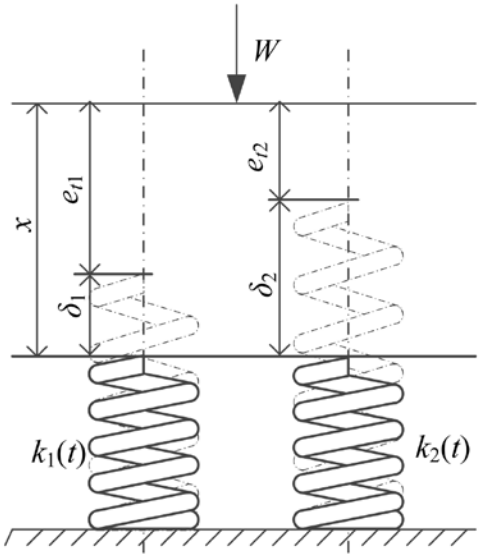


Fig. 3. Force diagram of meshing gears (condition 2)

$$\beta = \frac{W_1}{W} = \frac{k_1(t)}{k_1(t)+k_2(t)} + \frac{k_1(t)k_2(t)}{k_1(t)+k_2(t)}\frac{|\Delta|}{W} \tag{4}$$

Where $\Delta=e_{t1}-e_{t2}$ is geometric clearance while meshing. Because of $\Delta<0$, the clearance is occurred in the second gear pair.

The distribution load between the two gear pairs are deduced respectively as:

$$\begin{cases} W_1 = \dfrac{k_1(t)(W+|\Delta|\cdot k_2(t))}{k_1(t)+k_2(t)} \\ W_2 = \dfrac{k_2(t)(W-|\Delta|\cdot k_1(t))}{k_1(t)+k_2(t)} \end{cases} \tag{5}$$

As shown in Figure 3, the second gear pair mesh firstly, the clearance is occurred in the first gear pair because Δ is more than 0. Load distribution coefficient β is:

$$\beta = \frac{W_1}{W} = \frac{k_1(t)}{k_1(t)+k_2(t)} - \frac{k_1(t)k_2(t)}{k_1(t)+k_2(t)}\frac{|\Delta|}{W} \tag{6}$$

The distribution load between the two gear pairs are deduced respectively as:

$$\begin{cases} W_1 = \dfrac{k_1(t)(W-|\Delta|\cdot k_2(t))}{k_1(t)+k_2(t)} \\ W_2 = \dfrac{k_2(t)(W+|\Delta|\cdot k_1(t))}{k_1(t)+k_2(t)} \end{cases} \tag{7}$$

Considering the clearance caused by gear wear, mathematical expression of clearance Δ can be deduced as follows:

When the meshing point is at section B_1C, clearance Δ is:

$$|\Delta| = |h_1(y)+h_2(y)-(h_1(y+p_b)+h_2(y+p_b))| \tag{8}$$

When the meshing point is at section DB_2, clearance Δ is:

$$|\Delta| = |h_1(y)+h_2(y)-(h_1(y-p_b)+h_2(y-p_b))| \tag{9}$$

Where y is the distance between contact point i and pitch point i' along the direction of meshing line. $h(y)$ is wear depth per unit width.

3. A numerical prediction model of wear in spur gear

Under high contact pressure, relative rolling and sliding motion occurs at the surfaces of meshing teeth. Most gear sets have oil and grease lubrication; however, lubrication condition is boundary or mixed lubrication. It indicates that meshing surface can't be fully separated by lubricant and occurs directly metal rubbing. Thus, material transfer and peeling metal will be generated at teeth surface under the low speed and heavy-duty condition, which is called as gear wear.

A generalized wear equation, which is called Archard model, is usually used to predict gear wear:

$$\frac{V}{s} = K\frac{W}{H} \tag{10}$$

Where V is wear volume loss of material, s is sliding distance of meshing gear teeth, W is the applied normal load, H is material hardness of observed surface. K is dimensionless wear coefficient, which is related to lubrication condition and wear mechanism.

Flodin [9] proposed a mild wear model in spur gear. Supposing k is constant, and surface pressure p_i and sliding velocity v_i remain unchanged in a very short time, the wear depth can be expressed as:

$$h_{i,n} = h_{i,n-1} + \Delta t k N \sum_{j=1}^{s} p_{i,j} v_{i,j} \tag{11}$$

Where $h_{i,n-1}$ is wear depth of mesh point i after n-1 wear cycles, n is present wear cycles, Δt is time increment, N is running revolutions at every interval, during which the profile of gear teeth will not be updated. j is the point within Hertzian contact radius, $p_{i,j}$ is surface pressure of point j, $v_{i,j}$ is sliding velocity of point j.

Surface pressures of contact point i and j within contact radius can be calculated by the Winkler model [12, 17].

$$p_{i,j}(x_{i,j}) = \frac{(0.59\mathrm{E}^*)}{Ra}(a^2 - x_{i,j}{}^2) \tag{12}$$

Where, j is the meshing point, a is contact radius, $x_{i,j}$ is distance from point i in contact radius to meshing point j, E* is equivalent elastic modulus, R is equivalent radius of two rotating cylinders. The computation expressions of indirect variable in equation (12) are:

$$\frac{1}{E^*} = \frac{1-\mu_1^2}{E_1} + \frac{1-\mu_2^2}{E_2} \tag{13}$$

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} \tag{14}$$

$$R_1 = ia_1 = R_{b_1} \tan\beta_{1i} \tag{15}$$

$$R_2 = ia_2 = R_{b_2} \tan\beta_{2i} \tag{16}$$

$$a = \left(\frac{4WR}{\pi E^*}\right)^{1/2} \tag{17}$$

Where W represents W_1 or W_2, W_1 is shared load by the first teeth pair, W_2 is shared load by the second teeth pair, which can be obtained by equations (5) and (7), E_1, E_2, μ_1, μ_2 are respectively elastic modulus and poisson ratio of pinion and gear, R_{b1}, R_{b2} are radiuses of base circle, β_{1i}, β_{2i} are pressure angle of the mesh point i, Φ is engagement angle.

The sliding velocity v_i of meshing point i can be determined by gear mesh theory:

$$v_i = (\omega_1 + \omega_2) \times y_i \tag{18}$$

$$y_i = ii' = R_{b1}(\tan\beta_{1i} - \tan\phi) = R_{b2}(\tan\phi - \tan\beta_{2i}) \tag{19}$$

The flow chart for wear calculation is provide in Figure 4.

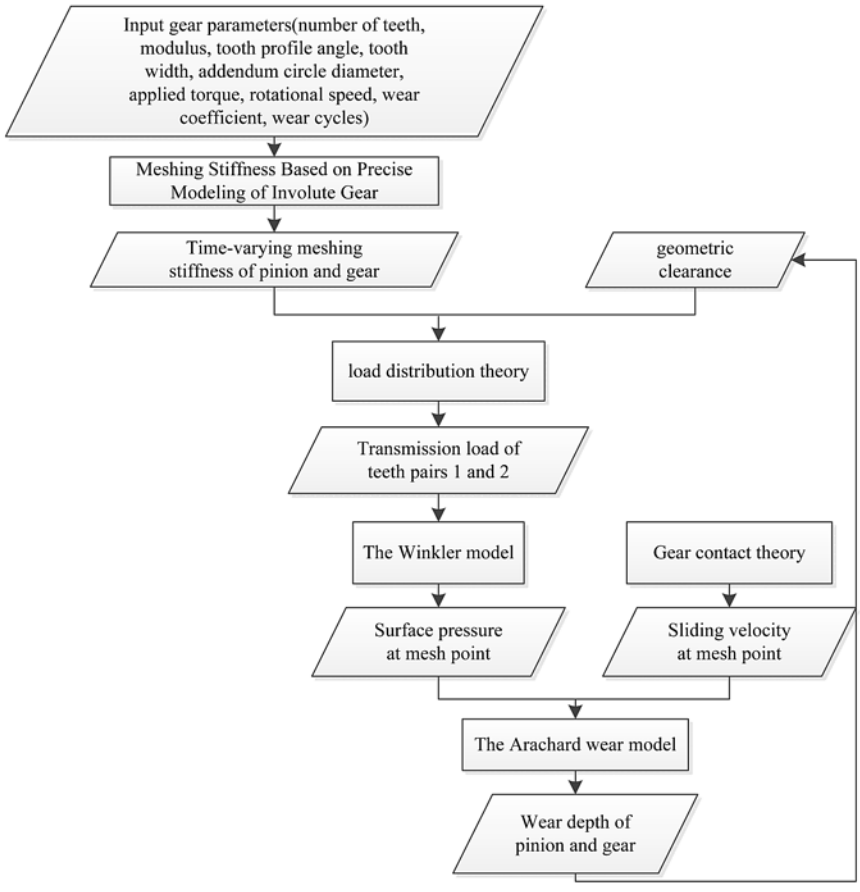


Fig. 4. The calculation flow chart

4. Surrogate models for nonlinear response

4.1. Artificial neural network

Because of the capabilities in pattern classification and function approximation, artificial neural network models are being one of most mature and widely used network models. A simple three-layer BP network can approach any complicated nonlinear function relations. A random vector $\mathbf{X}^{\mathrm{T}} = \{x_1, x_2, ..., x_n, t\}^{\mathrm{T}}$ which consists of original stochastic variables $\{x_1, x_2, ..., x_n\}^{\mathrm{T}}$ and time t is treated as the input parameter of network, and wear depth response $h_i(x_1, x_2, ..., x_n, t)$ is treated as the output parameter of network. The expression of nonlinear function is obtained after training:

$$h_i = f\left(\theta_k + \sum_{j=1}^{m} w_{kj}\delta\left(\theta_j + \sum_{i=1}^{n+1} w_{ji}x_i\right)\right) \tag{20}$$

where δ is S-type logarithmic function, f is linear function.

$$\delta(x) = \frac{1}{1+e^{-\beta x}}, \beta > 0 \tag{21}$$

$$f(x) = x \tag{22}$$

4.2. Kriging model

On the basis of the kriging method, the stochastic response model is denoted as two parts. The first part is the parametric regression model $F(\beta), x$. The second part is the non-parametric stochastic process model $z(x)$. The response is defined as:

$$h(x) = F(\beta, x) + z(x) = f(x)^{\mathrm{T}}\beta + z(x) \tag{23}$$

Where x is random variable and h is stochastic response, both of them are assumed to satisfy the normalization conditions. β is regression parameter. $f(x)$ is polynomial equation of x. $z(x)$ is random process and assumed to have mean zero and covariance is:

$$\mathrm{cov}\left[z(x_i), z(x_j)\right] = \sigma^2 R(\theta_k, x_i, x_j) \tag{24}$$

Where R is correlation function with 7 common forms, θ_k is the key parameter of the function.

To keep the predictor unbiased, The Kriging approximations are defined as:

$$\hat{\boldsymbol{h}}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^T \boldsymbol{\beta}^* + \boldsymbol{r}(\boldsymbol{x})^T \boldsymbol{R}^{-1}(\boldsymbol{H} - \boldsymbol{F}\boldsymbol{\beta}^*) \tag{25}$$

Where $\boldsymbol{x}=\{x_1, x_2, ..., x_{n+1}\}$ is a vector composed by untried point, n+1 is the number of random variables. $\boldsymbol{f}(\boldsymbol{x})^{\mathrm{T}}=\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), ..., f_p(\boldsymbol{x})\}$ are polynomials of vector $\boldsymbol{x}$, p is polynomial order. $\boldsymbol{\beta}^*=\{\beta_1^*, \beta_2^*, ..., \beta_p^*\}^{\mathrm{T}}$ is the generalized least squares solution. Given a set of m design points $\boldsymbol{x}_{s_1}, \boldsymbol{x}_{s_2}, ..., \boldsymbol{x}_{s_m}$, $\boldsymbol{r}(\boldsymbol{x})=\{R(\theta_k, \boldsymbol{x}, \boldsymbol{x}_{s_1}), R(\theta_k, \boldsymbol{x}, \boldsymbol{x}_{s_2}), ..., R(\theta_k, \boldsymbol{x}, \boldsymbol{x}_{s_m})\}^{\mathrm{T}}$ is the vector of correlation between untried points and design points. $R_{m\times m}$ is correlation matrix between design points. $\boldsymbol{H}$ is response vector of design points. $\boldsymbol{F}$ is design matrix of design points.

$\boldsymbol{\beta}^*$ is deduced according to the equation below:

$$\boldsymbol{\beta}^*=\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F}^{-1}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{H} \tag{26}$$

Where $R(\theta_k, \boldsymbol{x}_i, \boldsymbol{x}_j)$ is spatial correlation function of any two design points ($i, j \in s_1, ..., s_m$). And the generalized correlation function is Gaussian function.

$$R(\theta_k, \boldsymbol{x}_i, \boldsymbol{x}_j)=\prod_{k=1}^{n+1}\exp(-\theta_k \left| x_k^i - x_k^j \right|^2) \tag{27}$$

$R(\theta_k, \boldsymbol{x}_i, \boldsymbol{x}_j)$ is assumed to be assembled into the correlation matrix $\boldsymbol{R}$:

$$\boldsymbol{R} = \begin{bmatrix} \sigma_1^2 & R(x_1, x_2) & \dots & R(x_1, x_m) \\ R(x_2, x_1) & \sigma_2^2 & & \\ \dots & & & \\ R(x_m, x_1) & & & \sigma_m^2 \end{bmatrix}_{(m\times m)} \tag{28}$$

The optimal coefficients $\boldsymbol{\theta}^*$ of the correlation function is derived by maximum likelihood estimation method:

$$\boldsymbol{\theta}^* = \min_{\theta}\left\{ |\boldsymbol{R}|^{\frac{1}{m}} \sigma^2 \right\} \tag{29}$$

Where $|\boldsymbol{R}|$ is the determinant of $\boldsymbol{R}$.

If the form of correlation function, the order of polynomial and design samples are given, $\boldsymbol{\beta}^*$ can be calculated by equation (26). And $\hat{\boldsymbol{h}}(\boldsymbol{x})$ can be obtained by equation (25) when untried points $\boldsymbol{x}$ are given.

5. Reliability analysis method based on the Edgeworth series and four moment techniques

Gear wear accumulates gradually along with time. System failure is said to have occurred due to wear over the threshold value. As the wear speed and threshold of driving wheel are different from driven wheel, the reliability model of the driving and driven wheel need to be built respectively. Taking driving wheel as an example, according to the safety criterion that maximum wear depth on the tooth profile are not allowed to exceed specified clearance, the limited state function of driving wheel is defined as:

$$g_{\mathrm{p}}(x,t) = E_{\mathrm{sni}}^{\mathrm{p}} - \max(h_i^{\mathrm{p}}(x,t)) \tag{30}$$

Where $h_i(x,t)$ is wear depth of meshing point i at any time t, E_{sni} is lower deviation of tooth thickness and maximum clearance allowed.

The failure probability of driving wheel is formulated as follows:

$$P_{fc}^{\mathrm{p}}=\mathrm{P}\left[g_{\mathrm{p}}(x,t) \le 0 \right]=\mathrm{F}_{g_{\mathrm{p}}}(0) \tag{31}$$

The reliability of driving wheel can be expressed by equation (32):

$$R_{\mathrm{p}}(t)=\mathrm{P}\left[g_{\mathrm{p}}(x,t) > 0 \right] = 1 - P_{fc}^{\mathrm{p}} = \int_{g(x)>0} f(x)\mathrm{d}x \tag{32}$$

According to equation (32), the expression of limited state function and probability distribution of original variables are necessary precondition to calculate reliability. However, both of them can't be obtained due to lack of the statistical data and complex structure. To solve the problem, the arbitrary distribution function of the standard random variable is approximately expressed by the standard normal distribution function using the Edgeworth series[25] as follows:

$$F(y) = \Phi(y) - \varphi(y)\left[\frac{1}{6}\frac{\theta_g}{\sigma_g^3}H_2(y) + \frac{1}{24}\left(\frac{\eta_g}{\sigma_g^4} - 3 \right)H_3(y) + \frac{1}{72}\left(\frac{\theta_g}{\sigma_g^3} \right)^2 H_5(y) + \dots \right] \tag{33}$$

where $\phi(\cdot)$ is the standard normal probability density function, and $\mathrm{H}_i(y)$ is the Hermite polynomial, σ_g θ_g η_g are the variance, the third moment, and the fourth moment of the limited state function, which can be expressed by the corresponding moments of the random variables as follows based on the four moment method:

$$\mu_g = \mathrm{E}[g_d(\boldsymbol{X}) + \varepsilon g_p(\boldsymbol{X})] = g_d(\boldsymbol{X}) \tag{34}$$

$$\sigma_g{}^2 = \mathrm{Var}[g(\boldsymbol{X})] = \left(\frac{\partial g_d(\boldsymbol{X})}{\partial \boldsymbol{X}^{\mathrm{T}}} \right)^{[2]} \mathrm{Var}(\boldsymbol{X}) \tag{35}$$

$$\theta_{\mathbf{g}} = \mathrm{E}[\mathbf{g}(\boldsymbol{X}) - \mathbf{g}_d(\boldsymbol{X})]^3 = \left(\frac{\partial g_d(\boldsymbol{X})}{\partial \boldsymbol{X}^{\mathrm{T}}} \right)^{[3]} C_3(\boldsymbol{X}) \tag{36}$$

$$\eta_{\mathbf{g}} = \mathrm{E}[\mathbf{g}(\boldsymbol{X}) - \mathbf{g}_d(\boldsymbol{X})]^4 = \left(\frac{\partial g_d(\boldsymbol{X})}{\partial \boldsymbol{X}^{\mathrm{T}}} \right)^{[4]} C_4(\boldsymbol{X}) \tag{37}$$

Where $\mathrm{Var}(\boldsymbol{X})$ is the variance matrix that includes all variances and covariances of the random parameters. $\mathrm{C}_3(\boldsymbol{X})$, $\mathrm{C}_4(\boldsymbol{X})$ are the third and fourth moment matrices comprised by all the third and fourth central moments of each random parameter. $\mathrm{Var}(\boldsymbol{X})$, $\mathrm{C}_3(\boldsymbol{X})$, $\mathrm{C}_4(\boldsymbol{X})$ can be figured out by parameter estimation method. The Kronecker power is $(\cdot)^{[k]}=(\cdot)^{[k-1]}\otimes(\cdot)=(\cdot)\otimes(\cdot)\otimes\cdots\otimes(\cdot)$, and the symbol $\otimes$ represents the Kronecker product which is defined as $(\boldsymbol{A})_{\mathrm{p}\times\mathrm{q}}\otimes(\boldsymbol{B})_{\mathrm{s}\times\mathrm{t}}=[a_{ij}\boldsymbol{B}]_{ps\times qt}$. $\partial\mathbf{g}_d / \partial x_i$ can be calculated by taking a derivative with respect to ANN surrogate model.

Thus, the reliability R is represented as:

$$R_{\mathrm{p}}(t) = 1 - P_{fc}^{\mathrm{p}} = 1\text{-}\mathrm{F}(-\beta) \tag{38}$$

Where $\beta = \dfrac{\mu_g}{\sigma_g}$ is reliability index.

Substituting β into equation (33), the reliability of driving wheel is derived from equation (38).

6. Examples

Supposing the level of gear precision is 6, gear parameters are listed out in Table 1. The former is assumed to be deterministic, the latter are random and the four moments are given.

Table 1. Parameters of gear pair

Parameter	Code and size	Parameter	Code and size
Teeth number-driving	$z_1 = 24$	Modulus	m_{n}=4.5mm
Teeth number-driven	$z_2 = 26$	Pitch diameter-driving	d_{w1}=108mm
Center distance	a_{w}=112.5mm	Pitch diameter-driven	d_{w2}=117mm
Parameter	Mean value	The four moments	
Nominal pressure angle	$\alpha_{\mathrm{n}} = 20^{\circ}$	（20, 1, -0.00177972, 2.99779）	
Tooth width	$b = 15$mm	（15, 0.2025, 0.0081733, 0.123569）	
Applied torque	T_1=302000N · mm	（3.02e5, 9.1204e6, -1.35432e7, 2.4919e14）	
Driving gear speed	$n_1 = 150$r/min	(150, 100, 199.505, 30749.6)	
Wear coefficient	$k = 5\times10^{-10}$mm^{2}N^{-1}	(5e-10, 3.6e-21, -6.28011e-35, 3.88836e-41)	

6.1. Numerical solution of wear depth in gears

The time varying meshing stiffness can be calculated by the precise modeling of involute gear[23,24], as shown in Figure 5, the stiffness curve of a pair of teeth is given.

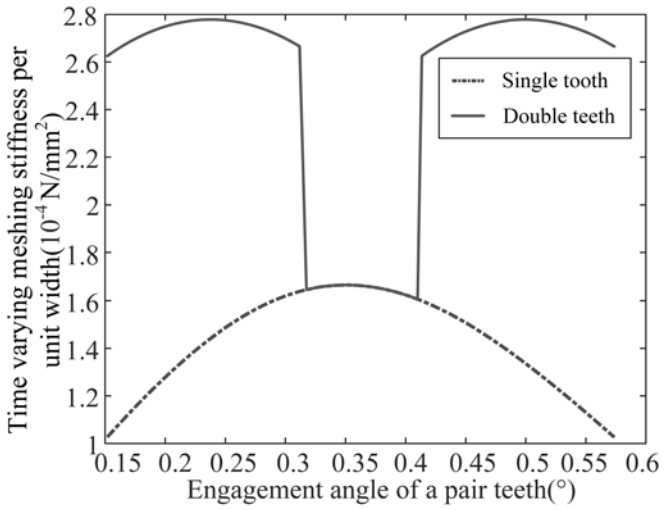


Fig. 5. Time varying meshing stiffness curves of single and double teeth based on precise modeling of involute gear

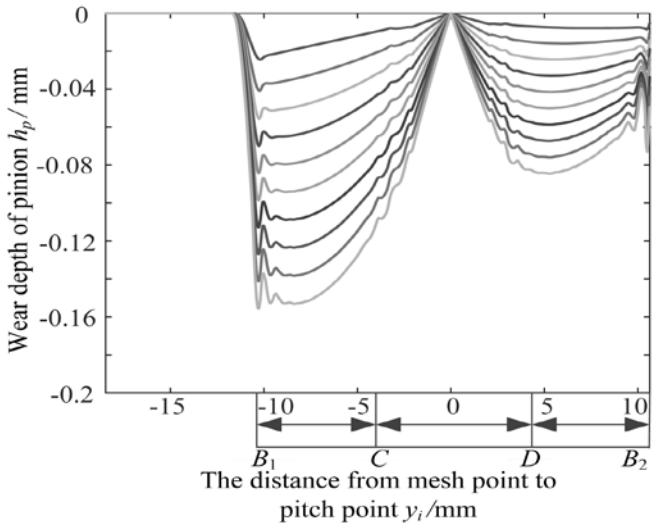


Fig. 6. Wear depth of the pinion after n wear cycles

Supposing the total wear cycle is 300, the gear wheel rotates 5000 circles (that is N=5000) after one wear cycle. In any circle, tooth profile, surface pressure and sliding velocity remain to be constant.

The wear of the pinion and gear after n wear cycles are presented in the figure (6) and (7). To facilitate making clear drawing, the results of wear depth after 30, 60, ……,300 cycles are plotted in figures by the distance from meshing point to pitch point on the horizontal axis. According to the principle of gear engagement shown in Figure 1, the B1 point is start of engagement and B2 is end of engagement. The zero value is at pitch point, the negative is at points from base circle to pitch point, and the positive is from pitch point to addendum circle. The wear depths of the pinion and gear respectively are varying over the teeth flanks with the maximum wear at the root and minimum wear at the pitch. The wear of driving pinion is more than driven gear, which agrees with the conclusion in references[1,21].

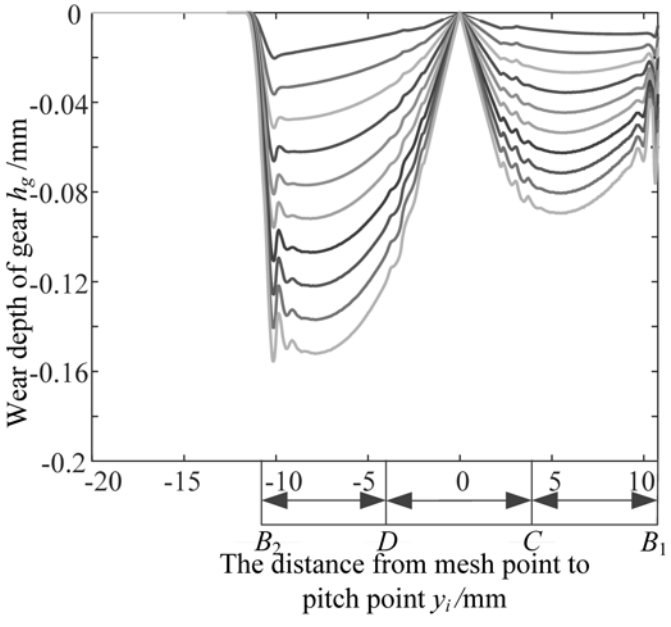


Fig. 7. Wear depth of the gear after n wear cycles

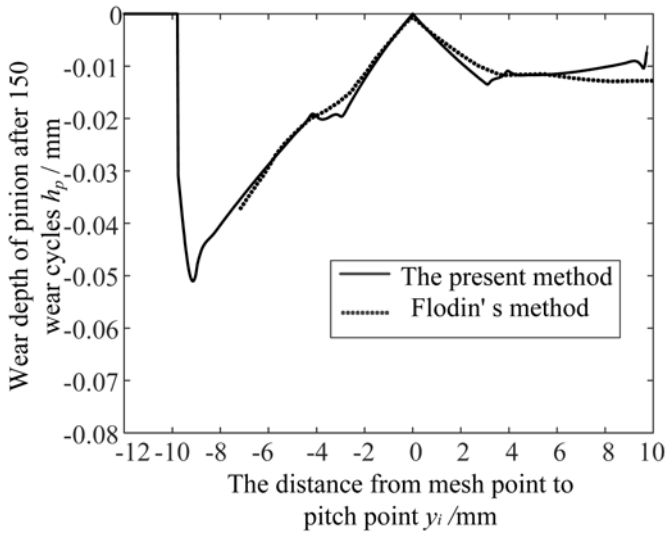


Fig. 8. Comparison of wear depth with the present method and Flodin's method

As the time-varying stiffness and the distribution load are both considered in present paper, wear depths at the conversion position of single teeth-meshing area and double teeth-meshing area are fluctuating slightly and don't appear sudden change while this phenomenon has appeared in reference[2] where the load is distributed equally. The wears are varying from pitch to addendum but drastically from root to pitch.

Based on the gear parameters in reference [9], wear depth with the method in the present paper are compared with Foldin's. Considering distributed load between two gear pairs, wear depths of the pinion with the two methods are shown in figure 8, and the results are very close.

6.2. ANN and Kriging surrogate models

A random vector $\boldsymbol{X}^{\mathrm{T}}=\{x_1,x_2,...,x_n,t\}^{\mathrm{T}}$ which consists of original stochastic variables $\{x_1,x_2,...,x_n\}^{\mathrm{T}}$ and time t is treated as the input parameter of network, and wear depth response $h_i(x_1,x_2,...,x_n,t)$ is treated as the output parameter of network. Three-layer BP neutral network models with a 6-13-1 form are structured when sample sizes are 50,100 and 500 respectively (see Appendix A). Log-sigmoid function is selected as the transfer function of hidden layer. Purelin function is selected as the transfer function of output layer. Trainlm func-

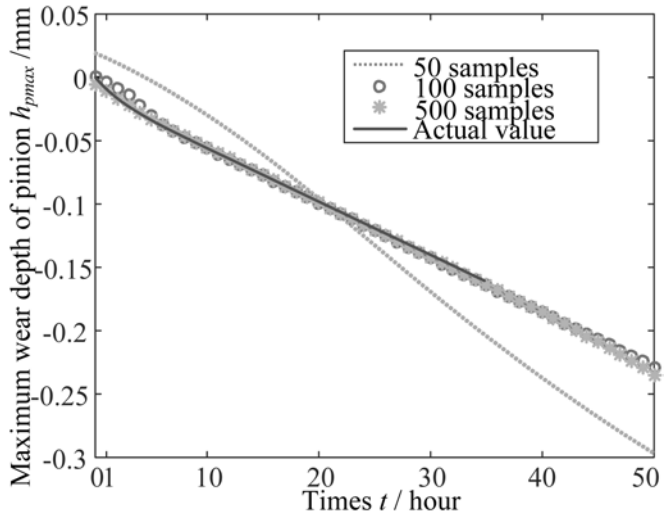


Fig. 9. Comparison of gradual wear process in driving wheel under different sample sizes with ANN method

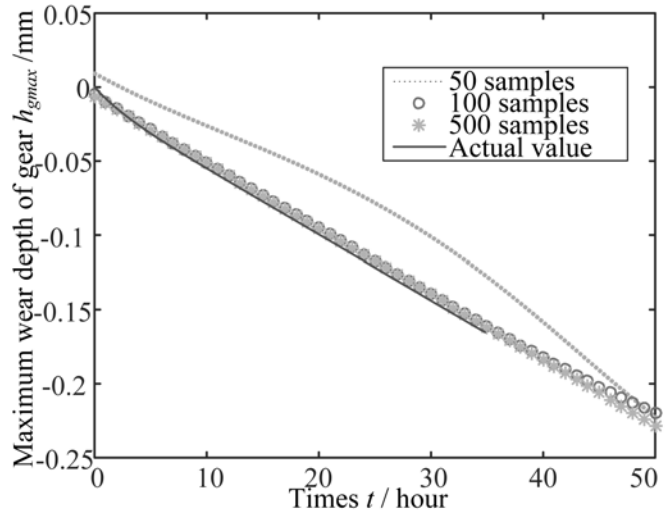


Fig. 10. Comparison of gradual wear process in driven wheel under different sample sizes with ANN method

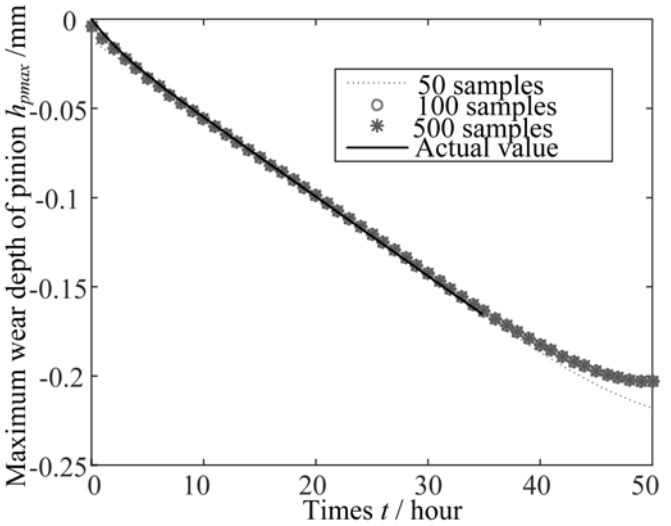


Fig. 11. Comparison of gradual wear process in driving wheel under different sample sizes with Kriging method

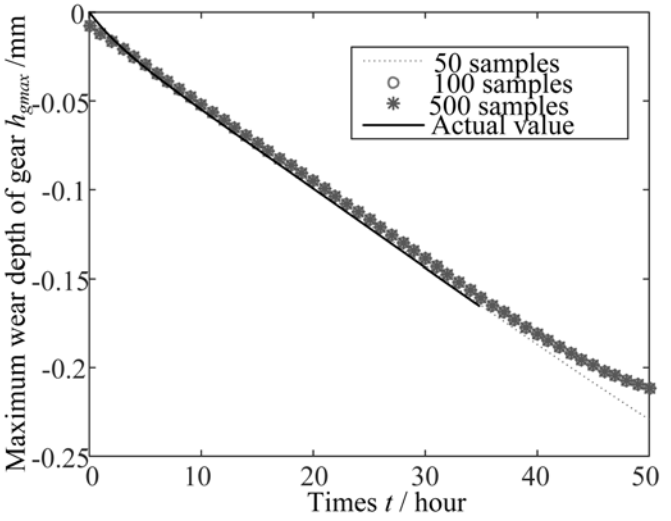


Fig. 12. Comparison of gradual wear process in driven wheel under different sample sizes with Kriging method

tion is selected as training function. The training terminates if network error is less than 10^{-6}. The training parameters of ANN are listed out in Appendix B.

To determine the suitable sample size, both of the accuracy and goodness of fit needed to be tested for the three network models with different samples. The accuracy means comparison of gradual wear process between exact and surrogate models when original stochastic variables $\{x_1, x_2, ..., x_n\}^T$ are deterministic and equal to the mean values. The goodness of fit means comparison of wear depths between the two models with another set of samples when time is deterministic.

The accuracy of driving and driven wheels with ANN model are figured as follows:

As shown in figures 9 and 10, the accuracy is low with 50 samples whereas high with 100 and 500 samples. And the latter two make accurate prediction on the developing trend of wear depth.

The accuracy of driving and driven wheels with Kriging model are figured as follows:

As shown in figures 11 and 12, all of three models have high accuracy and make accurate prediction on the developing trend. It shows that Kriging method needs less samples than ANN method under the same accuracy.

Figures 13 and 14 shows goodness of fit based on ANN and Kriging surrogate models with 100 samples. Results of the other two are shown in Table 2.

The goodness of fit for ANN models with 100 and 500 samples can reach more than 0.9999. But the 50-samples model has a low goodness of fit so that can't be used. The goodness of fit for Kriging models with all samples can reach more than 0.96. In a comprehensive view, the suitable sample size for ANN model is more than 100, for Kriging more than 50.

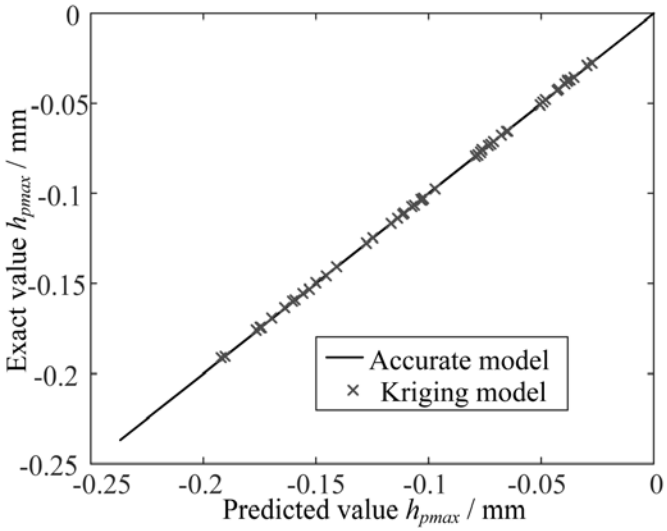


Fig. 13. Approximate values with Kriging method and Actual values of driving wheel

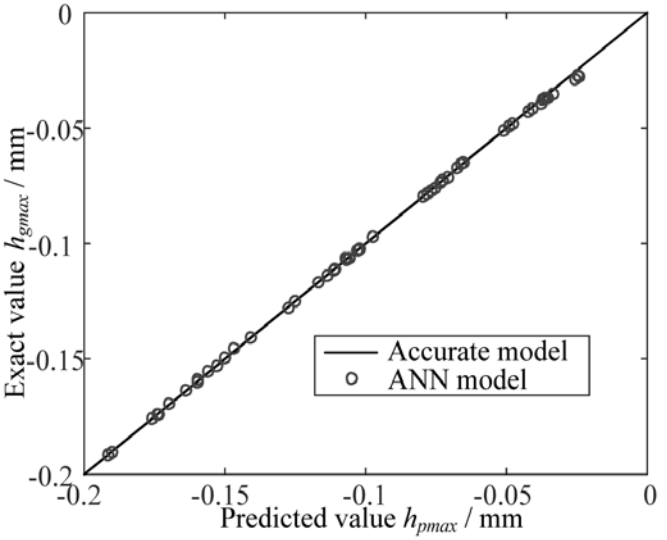


Fig. 14. Approximate values with Kriging method and Actual values of driven wheel

6.3. Reliability curve

The permitted maximum clearances of driving and driven wheels are -142μm and -143μm respectively. On the basis of ANN surrogate models, combining with Edgeworth series and four moment method, the time-dependent reliability curve can be obtained under the condition that the four moments of original variables are known but the probability distribution unknown. Reliability curve can be also obtained with Kriging-based MCS method. Supposing probability distribution are known is convenient to make comparison of the ANN-Edgeworth Series-Four moment method and Kriging-MCS method, which are shown in Figures 15 and 16.

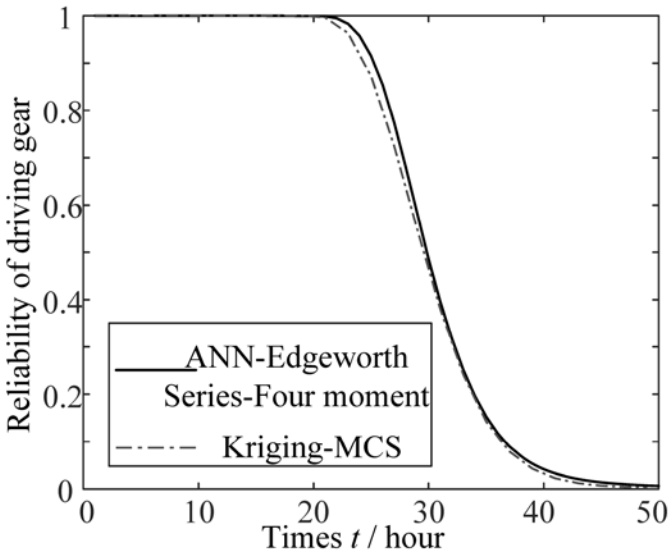


Fig. 15. Reliability curves of driving wheel

Table 2. Goodness of fit for ANN and Kriging models with different sample sizes

Sample size	ANN		Kriging	
	Driving wheel	Driven wheel	Driving wheel	Driven wheel
50	0.74318433	0.60040412	0.98699443	0.96489048
100	0.99990496	0.99997755	0.99997951	0.97556336
500	0.99997285	0.99997416	0.99997951	0.97556336

Reliability remains to be 1 in the beginning for a long time. After 20 hours, reliability decreases gradually and becomes 0 at 50 hours. It is important to note that the applied load in the present paper is much larger than that in practice to save time. And product life under normal load are longer than test load. However, the

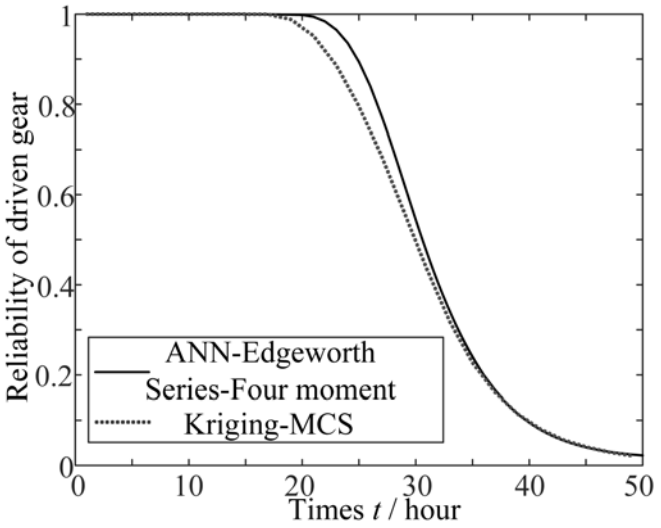


Fig.16. Reliability curves of driven wheel

Table 3. Consuming time of the three methods

Method	Sample size	Original samples	Program	Total
ANN-Edgeworth -Four moment	100	33.3 hours	10 seconds	34 hours
Kriging-MCS	50	16.7 hours	2 minutes	17 hours
Direct-MCS	10^5	3.8 years	10 seconds	3.8 years

present paper is focus on studying gradually law of wear reliability and choosing the suitable method to solve complex problem. According to the results with two methods, ANN-Edgeworth Series-Four moment method is a little bigger than Kriging-MCS method. That is to say, the difference is only a little. Supposing the time which is spent running the exact wear model once is 20 minutes, consuming time of every method for reliability is listed in Table 3.

The time for building up original samples account for absolute proportion, about more than 99% of the total. If more original samples the method needs, more time it will consume. As the direct-MCS method needs a large number of samples, the cost of time is as high as 3.8 years, which is impossible to complete. The former two only need tens of hours, which has a high efficiency.

7.conclusions

The present paper studied dynamic evolution law of mechanical reliability caused by wear. A numerical wear model has been established considering dynamic distribution load of gear tooth. Wear depths and gradual law has been obtained. Wear random process has been introduced. The problem of time-dependent reliability analysis with complex structural has been solved. The main conclusions are:

(1) Load and wear are coupled with each other. The dynamic characteristic of gear system is affected by variation of geometrical shape and dimension parameter caused by mild wear and dynamic load will lead to more serious wear. Thus, time-varying stiffness and dynamic distribution load should be both considered into wear model. As a result, sudden change hasn't appeared at the conversion point of single and double meshing area.

(2) Wear calculation spends a lot of time so that direct MCS method is not suitable for reliability analysis. It necessary to develop a surrogate model to replace exact model. Both of ANN and Kriging models have a high accuracy and goodness of fit. Kriging model needs less samples than ANN.

(3) Under the condition that the probability functions of random variables are unknown, Edgeworth series and four moment methods can be used to calculate reliability.

(4) Both of ANN-Edgeworth series-Four moment method and Kriging-MCS method are suitable for reliability with complex structure and have little difference.

Acknowledgment

We would like to express our appreciation to Chinese National Natural Science Foundation (51405072&U1708254), National Basic Research Program of China (973 Program) (2014CB046303).

Appendix A

Table 1. Original samples for construct ANN and Kriging models (50samples)

α_n	b	k	n_1	T_1	t	hgmax	hpmax
21.152	14.577	5.64E-10	144.85	294587.3	24853.3	-0.03447	-0.04364
21.152	14.577	5.64E-10	144.85	294587.3	49706.59	-0.06479	-0.07621
21.152	14.577	5.64E-10	144.85	294587.3	74559.89	-0.09522	-0.1085
21.152	14.577	5.64E-10	144.85	294587.3	99413.19	-0.12574	-0.14088
21.152	14.577	5.64E-10	144.85	294587.3	124266.5	-0.15607	-0.17364
19.778	14.986	4.69E-10	164.85	296417.6	21838.03	-0.03494	-0.03853
19.778	14.986	4.69E-10	164.85	296417.6	43676.07	-0.06283	-0.06556
19.778	14.986	4.69E-10	164.85	296417.6	65514.1	-0.09021	-0.09225
19.778	14.986	4.69E-10	164.85	296417.6	87352.14	-0.11774	-0.11871
19.778	14.986	4.69E-10	164.85	296417.6	109190.2	-0.14519	-0.14518
20.667	15.641	3.64E-10	155.15	298247.9	23203.35	-0.02374	-0.03
20.667	15.641	3.64E-10	155.15	298247.9	46406.7	-0.0434	-0.05069
20.667	15.641	3.64E-10	155.15	298247.9	69610.05	-0.06262	-0.07035
20.667	15.641	3.64E-10	155.15	298247.9	92813.41	-0.08188	-0.09002
20.667	15.641	3.64E-10	155.15	298247.9	116016.8	-0.10123	-0.10964
20.545	16.186	3.38E-10	124.85	300078.2	28834.6	-0.02206	-0.02766
20.545	16.186	3.38E-10	124.85	300078.2	57669.2	-0.04014	-0.04659
20.545	16.186	3.38E-10	124.85	300078.2	86503.8	-0.05755	-0.06439
20.545	16.186	3.38E-10	124.85	300078.2	115338.4	-0.07489	-0.08219
20.545	16.186	3.38E-10	124.85	300078.2	144173	-0.09221	-0.09982
19.899	13.977	3.89E-10	173.94	301908.5	20696.79	-0.03208	-0.03652
19.899	13.977	3.89E-10	173.94	301908.5	41393.58	-0.05761	-0.06161
19.899	13.977	3.89E-10	173.94	301908.5	62090.38	-0.08244	-0.0859
19.899	13.977	3.89E-10	173.94	301908.5	82787.17	-0.1073	-0.11022
19.899	13.977	3.89E-10	173.94	301908.5	103484	-0.1322	-0.13466
18.121	15.886	4.62E-10	162.42	303738.8	22164.76	-0.04406	-0.03855
18.121	15.886	4.62E-10	162.42	303738.8	44329.52	-0.07505	-0.06409
18.121	15.886	4.62E-10	162.42	303738.8	66494.27	-0.10521	-0.09326
18.121	15.886	4.62E-10	162.42	303738.8	88659.03	-0.13548	-0.12331
18.121	15.886	4.62E-10	162.42	303738.8	110823.8	-0.16565	-0.15342
18.323	14.55	4.15E-10	138.18	305569.1	26052.97	-0.04292	-0.03892
18.323	14.55	4.15E-10	138.18	305569.1	52105.95	-0.07331	-0.06432
18.323	14.55	4.15E-10	138.18	305569.1	78158.92	-0.10252	-0.09107
18.323	14.55	4.15E-10	138.18	305569.1	104211.9	-0.13183	-0.12003
18.323	14.55	4.15E-10	138.18	305569.1	130264.9	-0.16107	-0.14937
18.727	13.923	5.93E-10	152.12	307399.4	23665.53	-0.05548	-0.0523
18.727	13.923	5.93E-10	152.12	307399.4	47331.05	-0.0983	-0.09039
18.727	13.923	5.93E-10	152.12	307399.4	70996.58	-0.14104	-0.13121
18.727	13.923	5.93E-10	152.12	307399.4	94662.11	-0.18369	-0.17359
18.727	13.923	5.93E-10	152.12	307399.4	118327.6	-0.22698	-0.21605
20.263	14.277	3.42E-10	154.55	309229.7	23293.43	-0.02717	-0.03282
20.263	14.277	3.42E-10	154.55	309229.7	46586.87	-0.04913	-0.05515
20.263	14.277	3.42E-10	154.55	309229.7	69880.3	-0.07032	-0.07624
20.263	14.277	3.42E-10	154.55	309229.7	93173.73	-0.09143	-0.09723
20.263	14.277	3.42E-10	154.55	309229.7	116467.2	-0.11253	-0.11821
19.333	14.032	4.55E-10	143.64	311060	25062.66	-0.04084	-0.04273
19.333	14.032	4.55E-10	143.64	311060	50125.31	-0.07248	-0.07243
19.333	14.032	4.55E-10	143.64	311060	75187.97	-0.10338	-0.10184
19.333	14.032	4.55E-10	143.64	311060	100250.6	-0.13439	-0.13116
19.333	14.032	4.55E-10	143.64	311060	125313.3	-0.16523	-0.16045

Appendix B

Table 2. ANN network parameters of driving wheel

$[w_{ij}]$						$[w_{kj}]^T$	$[\theta_j]$	$[\theta_k]$
-5.22528	3.221899	24.04881	6.898731	2.740079	4.032272	-0.03077	-25.2424	
-0.06841	-0.65284	-1.66892	-3.01045	0.020653	-32.8203	0.108727	-26.8177	
-0.1639	1.876093	-0.32824	-0.93095	-0.80379	-0.14147	0.075344	-1.37843	
1.542938	-0.11529	-0.79372	-0.09672	0.060053	-3.12393	-0.09169	-0.76465	
0.06194	0.044609	-0.43121	-0.40289	-0.01041	-0.42879	4.683095	0.681912	
-0.28528	0.255703	1.007548	-0.78453	-0.03665	1.71228	-0.16621	-0.0539	
-0.05235	-0.10145	0.622529	0.442877	0.038716	-0.28308	2.116904	-0.67204	-50.9805
0.030045	0.100637	0.503045	-0.64295	0.015263	-0.97372	-1.26834	2.260346	
20.95529	-36.4756	-22.8803	3.913039	1.30388	5.377841	-0.00581	-16.6671	
0.193283	0.254279	-0.56163	-0.51449	-0.06677	-0.10976	73.87332	2.798588	
3.921268	-0.26835	-0.90545	-0.38699	0.114956	-1.64071	-25.2716	2.025285	
-0.18652	-0.23793	0.555175	0.509348	0.062009	0.032582	83.42451	-2.94615	
-3.9177	0.269415	0.902563	0.385194	-0.115	1.63199	-25.2959	-2.02856	

Table 3. ANN network parameters of driving wheel

$[w_{ij}]$						$[w_{kj}]^T$	$[\theta_j]$	$[\theta_k]$
10.01178	-17.8146	22.13516	1.454345	2.592017	5.142662	-0.01478	-25.1922	
0.11027	0.021028	0.363776	0.255979	-0.12823	-1.28498	-1.07677	2.467814	
-0.43166	-0.2898	0.638186	0.470036	-0.00758	0.002081	-4.09345	-2.03569	
-0.75759	0.216657	-0.46778	-1.69948	-0.11459	-31.0365	0.057017	-23.707	
0.044053	0.010625	-1.01128	0.65242	0.021075	0.0022	18.38031	0.173905	
-12.0028	8.778236	-23.4791	6.552863	13.23493	-4.59593	0.011259	20.71292	
-0.03192	-0.00961	1.015037	-0.65358	-0.01902	-0.02486	18.37038	-0.19951	-38.4905
0.143736	0.068286	-0.34145	-0.1845	-0.08919	-0.69346	4.502442	1.447148	
-3.04415	1.357087	-5.2511	1.515298	0.611239	-0.9471	0.021985	-0.02036	
-3.51331	-0.49121	1.019266	-0.42602	0.977701	1.087055	-0.04405	-2.40367	
-100.164	93.57081	-102.359	-49.2258	81.14516	3.081073	0.006283	-81.5959	
-0.31274	-0.21189	0.559268	0.363906	0.02836	-0.47226	7.562517	-2.52769	
2.749156	0.427162	-1.59628	-0.66692	-0.30248	-2.4033	17.67943	10.17811	

References

1. Amarnath M, Sujatha C, Swarnamani S. Experimental studies on the effects of reduction in gear tooth stiffness and lubricant filmthickness in a spur geared system. Tribology International 2009, 42: 340-352, https://doi.org/10.1016/j.triboint.2008.07.008.
2. Andersson S. Partial EHD theory and initial wear of gears, Stockholm: Royal Institute of Technology, 1975.
3. Archard J F. Contact and rubbing of flat surfaces. Journal of applied physics 1953, 24(8): 981-988, https://doi.org/10.1063/1.1721448.
4. Brandão J A, Martins R, Seabra J H O, Castro M J D. An approach to the simulation of concurrent gear micropitting and mild wear. Wear 2015, 324-325: 64-73, https://doi.org/10.1016/j.wear.2014.12.001.
5. Devooght J, Smidts C. Probabilistic reactor dynamics. I: The theory of continuous event trees. Nuclear science and engineering 1992, 111(3): 229-240, https://doi.org/10.13182/NSE92-A23937.
6. Dhanasekaran S, Gnanamoorthy R. Gear tooth wear in sintered spur gears under dry running conditions. Wear 2008, 265: 81-87, https://doi.org/10.1016/j.wear.2007.08.025.
7. Dugan J B, Bavuso S J, Boyd M A. Dynamic fault-tree models for fault-tolerant computer systems. Reliability, IEEE Transactions on 1992, 41(3): 363-377.
8. Flodin A. Wear of spur and helical gears, Stockholm: Royal Institute of Technology, 2000.
9. Flodin A, Andersson S. Simulation of mild wear in spur gears. Wear 1997, 207(1): 16-23, https://doi.org/10.1016/S0043-1648(96)07467-4.
10. Gomes H M, Awruch A M. Comparison of response surface and neural network with other methods for structural reliability analysis. structural safety 2004, 26(1): 49-67.
11. Hashim M, Yoshikawa H, Matsuoka T, Yang M. Considerations of uncertainties in evaluating dynamic reliability by GO-FLOW methodology - example study of reliability monitor for PWR safety system in the risk-monitor system. Journal of Nuclear Science and Technology 2013, 50(7): 695-708, https://doi.org/10.1080/00223131.2013.790304.

12. Johnson K L. CONTACT MECHANICS, Cambridge University Press, Cambridge, 1985.
13. Li X-T, Tao L-M, Jia M. A Bayesian networks approach for event tree time-dependency analysis on phased-mission system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2015; 17 (2): 273–281, http://dx.doi.org/10.17531/ein.2015.2.15.
14. Onishchenko V. Tooth wear modeling and prognostication parameters of engagement of spur gear power transmissions. Mechanism and Machine Theory 2008, 43: 1639-1664, https://doi.org/10.1016/j.mechmachtheory.2007.12.005.
15. Põdra P. FE Wear Simulation of Sliding Contacts, Stockholm: Royal Institute of Technology, 1997.
16. Park D, Kahraman A. A surface wear model for hypoid gear pairs. Wear 2009, 267(9): 1595-1604, https://doi.org/10.1016/j.wear.2009.06.017.
17. Po˜dra P, Andersson S r. Wear simulation with the Winkler surface model. Wear 1997, 207: 79-85, https://doi.org/10.1016/S0043-1648(96)07468-6.
18. Tan X H, Bi W H, Hou X L, Wang W. Reliability analysis using radial basis function networks and support vector machines. Computers and Geotechnics 2011, 38(2): 178-186, https://doi.org/10.1016/j.compgeo.2010.11.002.
19. Tombuyses B, Devooght J. Solving Markovian systems of O.D.E. for availability and reliability calculations. Reliability Engineering & System Safety 1995, 48: 47-55, https://doi.org/10.1016/0951-8320(94)00065-V.
20. Volovoi V. Modeling of system reliability Petri nets with aging tokens. Reliability Engineering & System Safety 2004, 84(2): 149-161, https://doi.org/10.1016/j.ress.2003.10.013.
21. Wang S R, Yan Y T, Ding J Y. Experimental Study on Mesh-wear of Involute Spur Gears. Journal of Northeastern University(Natural Science) 2004, 25(2): 146-149.
22. Wang Z, Xie L Y. Time-Dependent Reliability Theory and Method of Mechanical Component, Science Press, Beijing, 2012.
24. Zhang K X, Hu P, Zhang Y M. Numerical Calculation of Meshing Stiffness Based on Precise Modeling of Involute Gear. Machinery Design & Manufacture 2013, (2): 66-73.
25. Zhang Y M, Wen B C, Liu Q L. First passage of uncertain single degree-of-freedom nonlinear oscillators. Computer Methods in Applied Mechanics and Engineering 1998, 165(1-4): 223-231, https://doi.org/10.1016/S0045-7825(98)00042-5.
26. Zhang Y M, Wen B C, Liu Q L. Reliability sensitivity for rotor-stator systems with rubbing. Journal of Sound and Vibration 2003, 259(5): 1095-1107, https://doi.org/10.1006/jsvi.2002.5117.
27. Zhu Z F, Du X P. Reliability Analysis With Monte Carlo Simulation and Dependent Kriging Predictions. Journal of Mechanical Design 2016, 138(12): 121403, https://doi.org/10.1115/1.4034219.
28. Żurowski , Brzózka K, Górka B. Analysis of surface layers and wear products by Mössbauer spectral analysis, Wear 2013, 297: 958-965, https://doi.org/10.1016/j.wear.2012.10.012.

Lisha ZHU
1. School of Mechanical and Automotive Engineering
Zhaoqing University
Zhaoqing Avenue, Duanzhou District, Zhaoqing, Guangdong, 526061, P.R. China
2. School of Mechanical Engineering and Automation
Northeastern University
NO. 3-11, Wenhua Road, Heping District
Shenyang, Liaoning, 110819, P.R. China

Yimin ZHANG
1. School of Mechanical and Automotive Engineering
Zhaoqing University
Zhaoqing Avenue, Duanzhou District, Zhaoqing, Guangdong, 526061, P.R. China
2. Institute of Equipment Reliability
Shenyang University of Chemical Technology
NO. 11 Road, Economic and Technical Development Zone
Shenyang, Liaoning, 110142, P.R. China

Rui ZHANG
School of Mechanical Engineering and Automation
Northeastern University
NO. 3-11, Wenhua Road, Heping District
Shenyang, Liaoning, 110819, P.R. China

Peiming ZHANG
Guangxi Research Institute Mechanical Industry
NO. 5, Chuangxin Road, Xixiang tang District
Nanning, Guangxi, 530000, P.R. China

E-mails: neulisachu@163.com, ymzhang@mail.neu.edu.cn, 15940256305@163.com, 1357889106@qq.com

Article citation info:
ZHANG L, ZHANG J, ZHAI H, ZHOU S. A new assessment method of mechanism reliability based on chance measure under fuzzy and random uncertainties. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 219–228, http://dx.doi.org/10.17531/ein.2018.2.06.

Lei ZHANG
Jianguo ZHANG
Hao ZHAI
Shuang ZHOU

A NEW ASSESSMENT METHOD OF MECHANISM RELIABILITY BASED ON CHANCE MEASURE UNDER FUZZY AND RANDOM UNCERTAINTIES

NOWA METODA OCENY NIEZAWODNOŚCI MECHANIZMÓW OPARTA NA POMIARZE SZANSY WYSTĄPIENIA ZDARZENIA W WARUNKACH NIEPEWNOŚCI ROZMYTEJ I LOSOWEJ

The traditional reliability analysis methods based on probability theory and fuzzy set theory have been widely used in engineering practice. However, these methods are unable directly measure the uncertainty of mechanism reliability with uncertain variables, i.e., subjective random and fuzzy variables. In order to address this problem, a new quantification method for the mechanism reliability based on chance theory is presented to simultaneously satisfy the duality of randomness and the subadditivity of fuzziness in the reliability problem. Considering the fact that systems usually have multilevel performance and the components have multimode failures, this paper proposes a chance theory based multi-state performance reliability model. In the proposed method, the chance measure is adopted instead of probability and possibility measures to quantify the mechanism reliability for the subjective probability or fuzzy variables. The hybrid variables are utilized to represent the random and fuzzy parameters, based on which solutions are derived to analyze the chance theory based mechanism reliability with chance distributions. Since the input parameters of the model contain fuzziness and randomness simultaneously, an algorithm based on chance measure is designed. The experimental results on the case application demonstrate the validity of the proposed method.

***Keywords**: chance measure, reliability assessment, uncertainty quantification, mechanism reliability.*

Tradycyjne metody analizy niezawodności oparte na teorii prawdopodobieństwa i teorii zbiorów rozmytych znajdują szerokie zastosowanie w praktyce inżynierskiej. Jednak metod tych nie można stosować do bezpośredniego pomiaru niepewności niezawodności przy niepewnych zmiennych, tj. subiektywnych zmiennych losowych i rozmytych. Aby zaradzić temu problemowi, przedstawiono nową metodę kwantyfikacji niezawodności opartą na teorii szansy, która jednocześnie spełnia aksjomaty dwoistości losowości oraz subaddytywności związanej z rozmytością w problemach niezawodności. Biorąc pod uwagę fakt, że systemy zazwyczaj charakteryzują się wielopoziomową strukturą, a uszkodzenia elementów składowych mają charakter wieloprzyczynowy, w niniejszym artykule zaproponowano model niezawodności eksploatacji systemu wielostanowego oparty na teorii szansy. W proponowanej metodzie, zamiast miar prawdopodobieństwa i możliwości, do kwantyfikacji niezawodności, w przypadku gdy dane są subiektywne zmienne losowe lub zmienne rozmyte, przyjęto miarę szansy wystąpienia zdarzenia. Do reprezentacji parametrów losowych i rozmytych wykorzystano zmienne hybrydowe, które stanowią podstawę dla wyprowadzenia rozwiązań w celu analizy niezawodności mechanizmu opartej na teorii szansy z rozkładem szans. Ponieważ parametry wejściowe modelu noszą jednocześnie znamiona rozmytości i losowości, opracowano algorytm oparty na mierze szansy. Wyniki eksperymentalne otrzymane na podstawie studium przypadku dowodzą poprawności proponowanej metody.

***Słowa kluczowe**: miara szansy, ocena niezawodności, kwantyfikacja niepewności, niezawodność mechanizmu.*

1. Introduction

In practical engineering, various uncertainties are unavoidable due to the complicated environmental factors, incomplete knowledge and inevitable measurement errors [14]. Thus, the mechanism reliability analysis requires proper modeling of all sources of uncertainty. The reliability analysis based on probability theory and Boolean algebra has made many important achievements and has been widely used in engineering practice. The probabilistic reliability is considered as the most valuable issue in engineering. In probabilistic framework, the uncertainties are modeled as random variables or stochastic processes by using a large amount of sample statistical information [1,8]. The application of probabilistic reliability requires sufficient information to construct precise probability density functions of uncertain parameters, but the sample information is not always adequate in the early stage of numerical analysis and optimization design [2].

In practical engineering, besides the randomness that can be modeled by probabilistic theory with probability distribution functions, epistemic uncertainty is another issue, caused by factors such as loss of information, limited knowledge, and inevitable man-made mistakes [11] which cannot be well explained by randomness and probabilistic models. For uncertain problems in practical engineering, a random variable is always employed to represent a kind of subject probability that is conducted by experts' judgments (subjective interpretation) and the uncertainty of this variable is actually the fuzziness that comes from experts' judgments [29]. In order to overcome this shortcoming, Zadeh [36] developed fuzzy set theory in 1965. In 1975, Kaufmann [11] first used the fuzzy theory in reliability engineering. Up until

now, the fuzzy theory has received widespread attention for reliability problems with subjective uncertainties [3, 9]. In order to measure a fuzzy event, Zadeh [37] proposed the concept of possibility measure, and afterward many researchers introduced it into fuzzy reliability analysis [4, 32]. Although, the possibility measure has been widely used, it does not obey the law of truth conservation and is inconsistent with the law of excluded middle and the law of contradiction. The main reason is that the possibility measure has no self-duality property. However, a self-dual measure is definitely required in both theory and practice. In order to define a self-dual measure, Liu [15, 16] presented the concept of credibility measure in 2002, and constructed an axiomatic foundation for credibility theory. In addition, Li [17] provided a sufficient and necessary condition for credibility measure. From then on, the credibility theory became a branch of mathematics for studying the behavior of fuzzy phenomena. A survey of credibility theory can be found in Liu [18].

Fuzziness and randomness are the two basic types of uncertainties, and both may appear in a structural system simultaneously. A fuzzy probabilistic model was proposed by Holický [10] in 2006, which combined the two types of uncertainties in the newly defined fuzzy probabilistic measures of structural reliability, the damage function and the fuzzy probability of failure. In most practical situations, some input parameters of system might be represented with probability distribution functions, while some with membership functions [20]. For completeness, the different knowledge conditions for each uncertainty parameter derive "hybrid" uncertain variables in structural reliability analysis [33]. Therefore, randomness and fuzziness should be jointly considered to comprehensively and correctly analyze the reliability of systems, resulting in a hybrid model with random and fuzzy variables [5].

Numerous approaches have been proposed to solve the aforementioned hybrid reliability problems in systems. Most of these approaches separate random and fuzzy parameters based on a double-sampling framework [20]. In order to avoid the deterioration of efficiency and accuracy, several works have attempted to combine the stochastic expansions with traditional optimization methods [7, 30]. They are mainly concentrated on explaining the fuzzy variables by adopting the probability theory, and calculating the reliability based on a probability measure. However, the probability measure with additivity used by these methods fails to satisfy the subadditivity axiom of fuzziness, and the possibility measure cannot satisfy the duality axiom of randomness [11]. Therefore, a reliability quantification model based on probability theory and the one based on credibility theory frequently yield infeasible solutions with large differences and paradoxical results [14]. In other words, neither probability theory nor credibility theory can deal with mechanism reliability problems under epistemic uncertainty with hybrid subjective random and fuzzy variables, because the measures of the two theories cannot satisfy the duality and subadditivity simultaneously [11]. Thus, the development of a framework of hybrid reliability models that integrates the merits of different uncertainties is necessary.

In order to achieve a reasonable solution to these reliability problems, and solve the limitations of the two measures, the chance theory and the chance measure proposed by Liu [21] are introduced in mechanism reliability, including the normality, the duality, the subadditivity, and the product axioms. Chance theory is a hybrid of probability theory and credibility theory. This theory relies on the chance measure to describe the belief degrees of events affected by epistemic uncertainty. It provides a concrete mathematical description of different types of uncertain parameters in the chance space. The "chance measure" in the range of [0, 1] is adopted to represent the chance level about the occurrence of a particular event. Different from the chance theory used in this paper, Liu [28] combines the probability theory and the uncertain theory into a chance theory that also includes the normality, the duality, the subadditivity, and the product axioms. This theory relies on the chance measure to describe the belief degrees of events affected by human uncertainty and objective randomness. Uncertainty theory is a powerful tool for interpreting human uncertainty that was founded by Liu [19] and refined by Liu [26]. The study of uncertain random reliability analysis was started by Wen-Kang [35] with the concept of reliability index. They proposed a formula to calculate the reliability of a Boolean system involving both random and uncertain variables. However, in engineering practice, it is often necessary to evaluate the reliability of the structure in combination with the limit state function, and the parameters in the function may contain various types of uncertainty information (for instance about randomness and fuzziness). The reliability evaluation index based on the Boolean system is not suitable for this situation. Comparatively speaking, both of these chance theories not only have different theoretical foundation, but the types of uncertainty information considered by them are also different. In this paper, the mixture of fuzziness and randomness is mainly considered. Therefore, the chance theory based on the probability measure and the credibility measure is selected. Probability measure is used to deal with the parameters with sufficient information, while the credibility measure is employed to deal with the fuzzy variables. Moreover, Liu [18, 23] introduced a hybrid variable in 2006 as a tool to describe the quantities with fuzziness and randomness, and then proposed a general framework of hybrid programming. Based on the chance theory and the limit state function of structures, this paper explores a new quantification model, and applies it to quantify the performance reliability of structural systems with the hybrid uncertainty problem.

The remainder of this paper is organized as follows. In Section 2.1, some useful concepts in the credibility theory and the chance theory such as credibility measure, hybrid variable, and chance distribution are described; In Section 2.2, formulas based on the chance measure and the chance distributions are derived to quantify the uncertainty reliability and the uncertainty of failure with the performance state function of systems; According to the formulations based on the chance theory and the performance reliability theory, a chance theory based performance reliability model is defined in Section 3; An algorithm based on chance measure is designed in Section 4 followed by an engineering case presented in Section 5; Finally, conclusions are presented in Section 6.

2. Chance theory based quantification for mechanism reliability

2.1. Preliminaries of chance theory

Let Θ be a nonempty set with P as the power set of Θ. For any $A \in P(\Theta)$, Liu [24] presented a credibility measure $Cr\{A\}$ to express the chance that fuzzy event A occurs.

Definition 1 (Credibility measure [18]) The set function Cr is called a credibility measure if it satisfies the normality, monotonicity, self-duality, and maximality axioms.

Theorem 1 (Liu [24]) A fuzzy variable is a (measurable) function from a credibility space (Θ, P, Cr) to the set of real numbers.

Theorem 2 (Liu [24]) Let ξ be a fuzzy variable defined on the credibility space (Θ, P, Cr). Then its membership function is derived from the credibility measure by:

$$\mu(x) = (2Cr\{\xi = x\}) \wedge 1 \tag{1}$$

Theorem 3 (Credibility Inversion Theorem [24]) Let ξ be a fuzzy variable with membership function μ. Then for any set B of real numbers, we have:

$$Cr\{\xi \in B\} = \frac{1}{2}(\sup_{x\in B}\mu(x) + 1 - \sup_{x\in B^c}\mu(x)) \quad (2)$$

Theorem 4 (Product Credibility Axiom [18]) Let Θ_k be nonempty sets on which Cr_k ($k = 1,2,...,n$) satisfy the four axioms, respectively, and $\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_n$. Then, for each $(\theta_1,\theta_2,...,\theta_n) \in \Theta$

$$Cr\{(\theta_1,\theta_2,...,\theta_n)\} = Cr_1\{\theta_1\} \wedge Cr_2\{\theta_2\} \wedge \cdots \wedge Cr_n\{\theta_n\} \quad (3)$$

Theorem 5 (Credibility Subadditivity Theorem [16]) The credibility measure is subadditive. That is, for any events A, B

$$Cr\{A \cup B\} \le Cr\{A\} + Cr\{B\} \quad (4)$$

Chance measure was introduced by Liu [21] in 2009 as a tool to describe the quantities with fuzziness and randomness.

Definition 2 (Chance space [16]) Suppose that (Θ, P, Cr) is a credibility space and (Ω, A, Pr) is a probability space. Then, the product $(\Theta, P, Cr) \times (\Omega, A, Pr)$ is called a chance space.

Theorem 6 Let $(\Theta, P, Cr) \times (\Omega, A, Pr)$ be a chance space. A subset $\Lambda \subset \Theta \times \Omega$ is called an event if $\Lambda(\theta) \in A$ for each $\theta \in \Theta$, where $\Lambda(\theta) = \{\omega \in \Omega | (\theta,\omega) \in \Lambda\}$.

Definition 3 (Chance measure [21]) Let $(\Theta, P, Cr) \times (\Omega, A, Pr)$ be a chance space. Then a chance measure of event Λ is defined as:

$$Ch\{\Lambda\} = \begin{cases} \sup_{\theta\in\Theta}(Cr\{\theta\} \wedge \Pr\{\Lambda(\theta)\}), if & \sup_{\theta\in\Theta}(Cr\{\theta\} \wedge \Pr\{\Lambda(\theta)\}) < 0.5 \\ 1 - \sup_{\theta\in\Theta}(Cr\{\theta\} \wedge \Pr\{\Lambda^c(\theta)\}), if & \sup_{\theta\in\Theta}(Cr\{\theta\} \wedge \Pr\{\Lambda(\theta)\}) \ge 0.5 \end{cases} \quad (5)$$

Axioms 1 (Normality axiom). For the universal set $\Theta \times \Omega$, $Ch\{\Theta \times \Omega\} = 1$.

Axioms 2 (Duality axiom).For each event Λ, $Ch\{\Lambda\} + Ch\{\Lambda^c\} = 1$, where Λ^c is the complement set of Λ.

Axioms 3 (Subadditivity axiom).The chance measure is subadditive. That is, for any events Λ_1 and Λ_2

$$Ch\{\Lambda_1 \cup \Lambda_2\} \le Ch\{\Lambda_1\} + Ch\{\Lambda_2\} \quad (6)$$

Definition 4 (Chance distribution [21]) The chance distribution $\Phi : \Re \to [0,1]$ of a hybrid variable ξ is defined by:

$$\Phi(x) = Ch\{(\theta,\omega) \in \Theta \times \Omega | \xi(\theta,\omega) \le x\} \quad (7)$$

As two special hybrid variables, the chance distribution of a random variable ξ is just its probability distribution:

$$\Phi(x) = Ch\{\xi \le x\} = \Pr\{\xi \le x\} \quad (8)$$

and the chance distribution of a fuzzy variable ξ is just its uncertainty distribution:

$$\Phi(x) = Ch\{\xi \le x\} = Cr\{\xi \le x\} \quad (9)$$

In many cases, fuzziness and randomness simultaneously appear in a system. In order to describe these phenomena, Liu [18] introduced a hybrid variable as a measurable function from a chance space to the set of real numbers.

Definition 5 (Hybrid variable [18]) Let ξ be a measurable mapping function from a chance space $(\Theta, P, Cr) \times (\Omega, A, Pr)$ to the set of real numbers. Then, $\xi(\theta,\omega)$ is called a hybrid variable. If $\xi_1, \xi_2, ..., \xi_n$ are hybrid variables, and $f : \Re^n \to \Re$ is a measurable function, then $\xi = f(\xi_1, \xi_2, ..., \xi_n)$ is a hybrid variable defined as:

$$\xi(\theta,\omega) = f(\xi_1(\theta,\omega), \xi_2(\theta,\omega), ..., \xi_n(\theta,\omega)), \forall(\theta,\omega) \in \Theta \times \Omega \quad (10)$$

A hybrid variable $\xi(\theta,\omega)$ degenerates to a random variable when it does not vary with θ and to a fuzzy variable when it does not vary with ω. Fuzzy random variable [34] and random fuzzy [25] variable are instances of hybrid variable.

Example Let $\xi_1, \xi_2, \ldots, \xi_m$ and $\eta_1, \eta_2, \ldots, \eta_n$ be random and fuzzy variables, respectively. If f is a measurable function, then:

$$\tau = f(\xi_1, \xi_2, ..., \xi_m, \eta_1, \eta_2, ...\eta_n) \quad (11)$$

is a hybrid variable determined by:

$$\tau(\theta,\omega) = f(\xi_1(\omega), \xi_2(\omega), ..., \xi_m(\omega), \eta_1(\theta), \eta_2(\theta), ...\eta_n(\theta)) \quad (12)$$

for all $(\theta,\omega) \in (\Theta, P, Cr) \times (\Omega, A, Pr)$.

2.2. Chance theory based reliability

As generally known, reliability can typically be measured by the probability of structure functions that satisfy certain requirements. The structure functions can be expressed by the state function, which is determined by the failure criteria. Assume that $\mathbf{x} = (x_1, x_2, ..., x_n)$ are the n -dimensional input variables denoting the various factors that affect the structure functioning. Then, $G = g(x_1, x_2, ..., x_n)$ is the state function of systems, and $G = 0$ is the limit state function of variable space, which is also called the critical surface. The basic variable space can be divided into two parts, failure region and safe region, by the critical surface. Particularly, when the input variables $\mathbf{x} = (x_1, x_2, ..., x_n)$ consist of the random variables $\mathbf{x_R} = (x_{R1}, x_{R2}, ..., x_{Rn_r})$ and the fuzzy variables $\tilde{\mathbf{x}}_\mathbf{F} = (\tilde{x}_{F1}, \tilde{x}_{F2}, ..., \tilde{x}_{Fn_f})$ simultaneously, the state function of the structural system is a hybrid variable $G = g(\mathbf{x_R}, \tilde{\mathbf{x}}_\mathbf{F})$. With the state function of structural systems under epistemic uncertainty, this paper proposes a reliability definition based on the chance theory.

Definition 6 (Chance theory based reliability) Given a chance space $(\Theta, P, Cr) \times (\Omega, A, Pr)$, let the state function $G = g(\mathbf{x_R}, \tilde{\mathbf{x}}_\mathbf{F})$, where $G > 0$ and $G \le 0$ indicate safe state and failure state, respectively. Then the chance measure of occurrence of a failure event is defined as:

$$Ch_{failure} = Ch\{g(\mathbf{x_R}, \tilde{\mathbf{x}}_\mathbf{F}) \le 0\} \quad (13)$$

From the duality axiom of the chance theory, the chance theory based reliability of a structural system where $G = g(\mathbf{x_R}, \tilde{\mathbf{x}}_\mathbf{F}) > 0$ can be formulated as:

$$Ch_{reliability} = 1 - Ch_{failure} \quad (14)$$

Following the probability definition of reliability, $Ch_{reliability}$ denotes the chance theory based reliability to quantify the uncertainty of a safe event in a system with the numerical value of [0, 1], and $Ch_{failure}$ describes the chance of occurring a failure event. Due to the similarity with the probability definition of failure and reliability, the numerical value of Ch is used instead of frequency to represent the chance with which it is believed that the event will occur. The higher $Ch_{reliability}$ is the more chance the reliability event will happen.

If the chance distribution Φ is given, the failure uncertainty with Eq.(13) can be obtained by:

$$Ch_{failure} = Ch\{G \le 0\} = \Phi(0) \quad (15)$$

Meanwhile, the chance theory based reliability Eq. (14) transform into:

$$Ch_{reliability} = Ch\{G > 0\} = 1 - Ch\{G \le 0\} = 1 - \Phi(0) \quad (16)$$

3. Mechanism reliability analysis method under random and fuzzy uncertainties

3.1. Performance reliability assessment of multi-state system

In the real world, usually systems and their components perform their tasks at several levels of performance or exhibit multiple performance levels or states, and most systems gradually degrade and have a wide range of working efficiencies [13]. Thus, the state of the system may range from perfect functioning to complete failure in engineering practice. The fundamental of multi-state systems was introduced by Murchland [31] in the middle of 1970s. For a multi-state system, $\mathbf{S} = (S_0, S_2, ..., S_q)$ is the vector of function states of the system, it is assumed that there are *w* failure modes, the failure state space is $\bar{\Omega} = \{S_0, ..., S_{w-1}\}$, and $\Omega = \{S_w, ..., S_q\}$ is the state space that the system can work. The critical surface is given by the function $g_i(x_1, x_2, ..., x_n) = 0$ for failure mode $S_i (i = 0,1,..., w-1)$. It should be noted that for each function g_i, it is necessary that all $x_j (j = 1,2,...,n)$ are included in the function. Suppose that $Y = f_t(x_1, x_2, ..., x_n)$ represents a joint probability density function of performance variables at time t, then the probability that the structure will fail in mode S_i $(i = 0,1,..., w-1)$ up to time t is given by (for a larger-is-better case):

$$F_i(t) = \Pr(f_t(x_1, x_2, ..., x_n) \in S_i) = \Pr(g_i(x_1, x_2, ..., x_n) \le 0) \quad (17)$$

where S_i is the space determined by $g_i(x_1, x_2, ..., x_n) \le 0$, $i = 0,1,..., w-1$. The reliability considering only failure mode S_i is given by:

$$R_i(t) = 1 - F_i(t) \quad (18)$$

The overall mechanism reliability considering all *w* failure modes can be evaluated as:

$$R(t) = 1 - F(t) = 1 - \Pr(g_i(x_1, x_2, ..., x_n) \le 0, i = 0,1,..., w-1) \quad (19)$$

According to the above content, chance theory is combined with performance reliability theory and a definition of performance reliability is proposed as follows:

Definition 7 The chance based performance reliability is defined as the chance for the system (component) to function properly over a period of time t. In order to express this relationship mathematically, the variable T is defined to be the time to failure of the system (component), $T \ge 0$ .Then the performance reliability can be expressed as:

$$R(t) = 1 - \Phi(t) = Ch\{T \ge t\} \quad (20)$$

where $R(t) \ge 0$ and $R(0) = 1$.

From definition 7 and Eq.(13), the chance theory based performance reliability function can be represented as:

$$\begin{aligned} R\{Y\} &= Ch\{Y \in \Omega\} \\ &= Ch\{f_i(x_1, x_2, ..., x_n) \in \Omega, w \le i \le q\} \\ &= Ch\{g_i(x_1, x_2, ..., x_n) > 0, w \le i \le q\} \\ &= Ch_{reliability} \end{aligned} \quad (21)$$

where $\Omega = \{S_w, ..., S_q\}$ is the state space that the system can work.

When the performance variables $\mathbf{x} = (x_1, x_2, ..., x_n)$ consist of the random variables $\mathbf{x_R} = (x_{R1}, x_{R2}, ..., x_{Rn_r})$ and the fuzzy variables $\tilde{\mathbf{x}}_\mathbf{F} = (\tilde{x}_{F1}, \tilde{x}_{F2}, ..., \tilde{x}_{Fn_f})$ simultaneously, the randomness and fuzziness should be jointly considered. Therefore, a multi-state performance reliability model with random and fuzzy variables is proposed in the next section to comprehensively and correctly analyze the reliability of system.

3.2. Multi-state performance reliability model considering random and fuzzy variables

A system with two-state spaces is considered as an example, and the multi-state performance reliability model is established as shown in Fig.1.

According to the performance reliability model proposed in literature [6], a multi-state performance reliability model based on hybrid variables is proposed. In this section, the chance theory is applied to establish a new model for the performance reliability of the structural system in the presence of both random and fuzzy information. From the mathematical view, the conceptual system model is shown in Fig.2.

The symbols and the parameters are defined as: t is the operation time of the system; $\mathbf{Y} = (Y_1(t), ..., Y_n(t))$ is the vector of performance parameters of the system; $\mathbf{y} = \left[y_{ij}(t) \right]_{n \times q}$, $i = 1,2,...,n$, $j = 0,1,2,...,q$ denotes the criteria of the system; $\{Y(t) \in \Omega\}$ is the set of events that product's state is normal; and $R\{Y(t) \in \Omega\}$ is the performance reliability of the system.

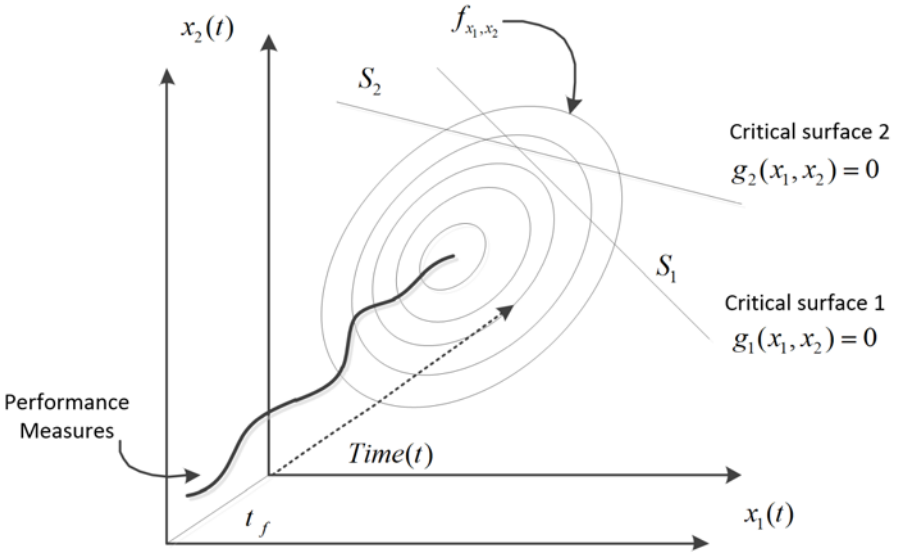


Fig. 1. The two-state performance reliability assessment with multiple input variables

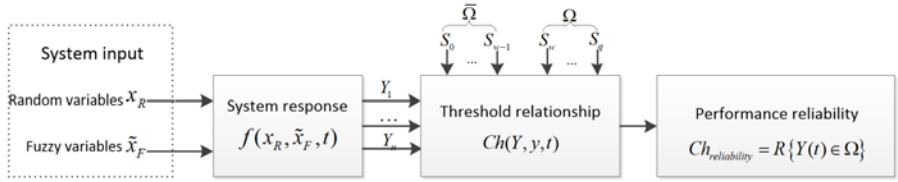


Fig. 2. Performance reliability model considering fuzzy and random factors

In Fig.2, $f(\mathbf{x_R},\tilde{\mathbf{x}}_\mathbf{F},t)$ is the state function of the system. From the definition 5 it can be seen that $f(\mathbf{x_R},\tilde{\mathbf{x}}_\mathbf{F},t)$ is a hybrid variable. The chance measure $Ch(\mathbf{Y},\mathbf{y},t)$ is a transforming operator that can transform inputs $\mathbf{Y}$ and $\mathbf{y}$ into $R\{Y(t)\in\Omega\}$. Thus, the mathematical model of the system is:

$$\begin{cases} Y(t)=f(x_R(t),\tilde{x}_F(t)) \\ Ch(Y(t),y(t))=R\{Y(t)\in\Omega\} \end{cases} \tag{22}$$

4. Simulation method of mechanism reliability model considering random and fuzzy variables

4.1. Statistical analysis model

In order to solve the general fuzzy programming models, Liu [23] proposed a fuzzy simulation algorithm based on credibility distribution and applied it to the solutions of uncertain functions and their expected functions. Later on, Chen [2] proposed a fuzzy based performance reliability algorithm, and designed a Monte Carlo method to estimate the fuzzy based model. According to these two methods, the following procedure may be used to handle the general simulation models considering random and fuzzy variables. θ_k is randomly generated from the credibility space (Θ,P,Cr), and ω_k the probability space (Ω,A,Pr), followed by writing $\nu_k=(2Cr\{\theta_k\})\wedge 1$ and producing $\tilde{x}_{Fk}=\xi(\theta_k)$, and $x_{Rk}=\omega_k$, $k=1,2,...,N$ respectively. Equivalently, $\tilde{x}_{Fk}$ and x_{Rk} are randomly generated and $\nu_k=\mu(\tilde{x}_{Fk})$ is written for $k=1,2,...,N$, where μ is the membership function of $\tilde{x}_F$.

Let the system simulate for N-times under the influence of random and fuzzy causes. The sampled values of characteristics parameters x_R and $\tilde{x}_F$ are $x_{R1r}, x_{R2r},\ldots, x_{Rn_r r}$ and $\tilde{x}_{F1r}, \tilde{x}_{F2r},\ldots,\tilde{x}_{Fn_f r}$ in the r-th simulation, respectively. From Eq.(23), the sampled values of performance parameters are:

$$\begin{cases} Y_{1r}=f_1(\mathbf{x_{Rr}},\tilde{\mathbf{x}}_\mathbf{Fr}) \\ \quad\ldots\ldots \\ Y_{nr}=f_n(\mathbf{x_{Rr}},\tilde{\mathbf{x}}_\mathbf{Fr}) \end{cases} \tag{23}$$

According to the model proposed in Section 3.2, the incident $\mathbf{S}=[S_0,S_1,...,S_q]$ is determined by the performance parameter $\mathbf{Y}=(Y_{1r}(t),...,Y_{nr}(t))$ and criterion $\mathbf{y}=\left[y_{ij}(t)\right]_{n\times q}$, $i=1,2,...,n$, $j=0,1,2,...,q$. The relationship is denoted by matrix:

$$\mathbf{Y}\times\mathbf{y}=\mathbf{S} \tag{24}$$

Eq.(24) is a logic expression that can be explained as follow: the incident S_0 occurs when the element of vector $(Y_{1r},Y_{2r},...,Y_{nr})$ meets the element of vector $(y_{10},y_{20},...,y_{n0})$. $S_1,S_2,...,S_q$ can be explained in a similar way. Alternatively, if one of the vectors $(Y_{1r},Y_{2r},...,Y_{nr})$ meets criterion y_j, and the others meet criterion y_{j+1}, the function state of the system is still S_j.

According to the results of N-times simulations, if the results of k-times simulations belong to Ω, then the reliability obtained by statistical analysis is $\frac{k}{N}$.

4.2. Chance theory based reliability simulation method

Assume that the structural system is reliable when the system state belongs to the set $\Omega=\{S_w,S_{w+1},...,S_q\}$, and is unreliable when the system state belongs to the set $\bar{\Omega}=\{S_1,S_2,...,S_{w-1}\}$. Thus, according to Eq.(21), the critical surface is given by the function $g_i(x_1,x_2,...,x_n)=0$ for state S_i, and the performance reliability of the system is:

$$\begin{aligned} Ch_{reliability} &= Ch\{Y\in\Omega\} \\ &= Ch\{f_j(x_R,\tilde{x}_F)\in\Omega, j=1,2,...,n\} \\ &= 1-Ch(g_i(x_R,\tilde{x}_F)\le 0, j=1,2,...,n) \end{aligned} \tag{25}$$

In order to compute the uncertain function $Ch\{g_j(x_R,\tilde{x}_F)\le 0, j=1,2,...,n\}$, using the method proposed in Section 4.1, $\tilde{x}_{F1},\tilde{x}_{F2},...,\tilde{x}_{FN}$ and $x_{R1},x_{R2},...,x_{RN}$ are randomly

generated from the credibility space (θ, P, Cr) and the probability space $(\Omega, A, \Pr)$, respectively. For each $\tilde{x}_{Fk}$,

$$\Pr\left\{g_j(x_R,\tilde{x}_{Fk})\le 0, for \quad all \quad j\right\}, \Pr\left\{g_j(x_R,\tilde{x}_{Fk})>0, for \quad some \quad j\right\} \quad (26)$$

is estimated by stochastic simulation via the samples $x_{R1}, x_{R2}, ..., x_{RN}$.

Besides, by using the credibility inversion theorem Eq.(2), the expression in Eq.(5) can be presented as follows:

$$\begin{aligned}&\max_{1\le k\le N} Cr\left\{\tilde{x}_{Fk}\right\}\wedge \Pr\left\{g_j(x_R,\tilde{x}_{Fk})\le 0, for \quad all \quad j\right\}\\ &= \max_{1\le k\le N}\left\{\frac{1}{2}(\sup_{\tilde{x}_{Fk}\in\Omega}\mu(\tilde{x}_{Fk})+1-\sup_{\tilde{x}_{Fk}\in\bar{\Omega}}\mu(\tilde{x}_{Fk}))\right\}\wedge \Pr\left\{g_j(x_R,\tilde{x}_{Fk})\le 0, for \quad all \quad j\right\}\\ &and\\ &\max_{1\le k\le N} Cr\left\{\tilde{x}_{Fk}\right\}\wedge \Pr\left\{g_j(x_R,\tilde{x}_{Fk})>0, for \quad some \quad j\right\}\\ &= \max_{1\le k\le N}\left\{\frac{1}{2}(\sup_{\tilde{x}_{Fk}\in\Omega}\mu(\tilde{x}_{Fk})+1-\sup_{\tilde{x}_{Fk}\in\bar{\Omega}}\mu(\tilde{x}_{Fk}))\right\}\wedge \Pr\left\{g_j(x_R,\tilde{x}_{Fk})>0, for \quad some \quad j\right\}\end{aligned} \quad (27)$$

Then according to Eq.(5), after simulating for N-times, if:

$$\max_{1\le k\le N} Cr\left\{\tilde{x}_{Fk}\right\}\wedge \Pr\left\{g_j(x_R,\tilde{x}_{Fk})\le 0, for \quad all \quad j\right\}<0.5 \quad (28)$$

then the value of $Ch\left\{g_j(x_R,\tilde{x}_F)\le 0, j=1,2,...,n\right\}$ is:

$$Ch\left\{g_j(x_R,\tilde{x}_F)\le 0\right\}=\max_{1\le k\le N} Cr\left\{\tilde{x}_{Fk}\right\}\wedge \Pr\left\{g_j(x_R,\tilde{x}_{Fk})\le 0, for \quad all \quad j\right\} \quad j=1,2,...,n \quad (29)$$

Otherwise, if:

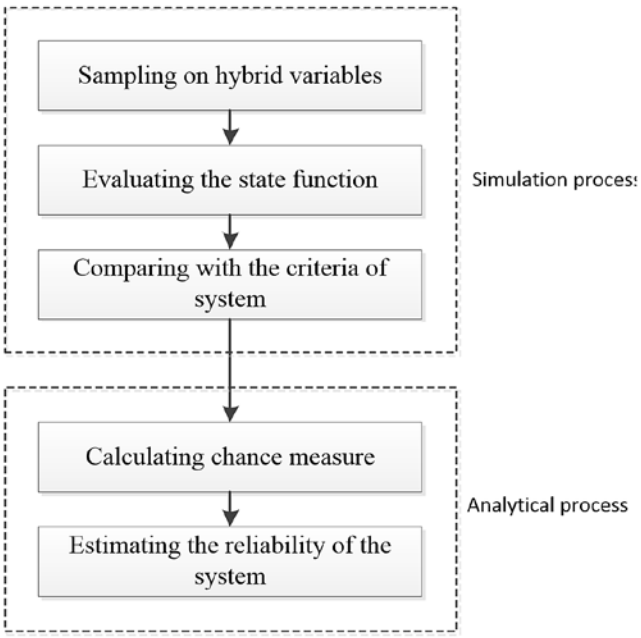


Fig. 3. The flowchart of the chance measure based simulation method

$$\max_{1\le k\le N} Cr\left\{\tilde{x}_{Fk}\right\}\wedge \Pr\left\{g_j(x_R,\tilde{x}_{Fk})\le 0, for \quad all \quad j\right\}\ge 0.5,$$

then:

$$Ch\left\{g_j(x_R,\tilde{x}_F)\le 0\right\}=1-\max_{1\le k\le N} Cr\left\{\tilde{x}_{Fk}\right\}\wedge \Pr\left\{g_j(x_R,\tilde{x}_{Fk})>0, for \quad some \quad j\right\} \quad j=1,2,...,n \quad (30)$$

Finally, the chance theory based reliability $Ch_{reliability}$ of the multi-state system can be calculated by Eq.(25).

Considering fuzzy and random variables, the chance theory based simulation method is proposed to conduct reliability calculation in Eq.(25). Five steps are involved in the proposed method. The first three steps are the simulation process, while the last two steps are the analytical process. The flowchart of the proposed method is shown in Fig.3.

5. An illustrated example

The main failure model of the harmonic gear reducer is due to the increase of the clearance between the components caused by the wear cumulating [27]. The clearance affects the transmission error $\Delta\varphi_{hg}$ and the backlash j_φ of the harmonic gear reducer, reducing the accuracy.

Fig.4 shows a harmonic gear reducer, the wave generator H is active, circular spline is fixed Z_R, and flex spline Z_G is the output.

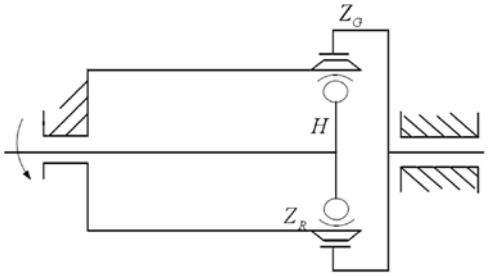


Fig. 4. Sketch map of the harmonic gear reducer

The speed of wave generator is $n=100r/\min$, load is $T=10N\cdot m$, the gear pressure angle is $\alpha=28.6^\circ$, the teeth number of the flex spline and the circular spline are $z_1=200$ and $z_2=202$, respectively and the harmonic gear ratio is $i=100$. When the transmission error and the backlash are greater than 0.15°, the harmonic gear reducer is considered to be invalid. Because of the high accuracy and lubrication conditions, the harmonic gear reducer can be considered to work in the stable wear stage. The wear rate is constant and the wear amount increases linearly with time. If the wear rate is μ, the cumulative wear is $W(t)=\mu t$. Therefore, the calculation formula of the transmission error $\Delta\varphi_{hg}$ of the harmonic gear reducer considering the amount of wear is:

$$\begin{aligned}\Delta\varphi_{hg}=\pm\frac{k_B}{\sqrt{N}}\Bigg[&0.2\times(\Delta F_{\sum j}+\frac{\pi d_1}{1.2Ui}(\Delta\rho_n+W(t)))+\\ &0.3\times\sqrt{(\Delta F_1)^2+(\Delta F_2)^2+(E_{rb})^2+(\frac{\pi d_1}{1.2Ui})^2[\sum_1^6(\Delta\rho_j)^2+W(t)^2]}\Bigg]\frac{412.8}{d_1}\end{aligned} \quad (31)$$

The backlash of the harmonic gear reducer j_φ is calculated as:

$$j_\varphi = \frac{6.876\tan\alpha(E_M + u_r + 2(\Delta f_a + W(t)) - 2\Delta F_r)}{mz_2} \quad (32)$$

According to the failure criterion, four states of the harmonic gear reducer can be defined:

$$S_0 : \Delta\varphi_{hg} < 0.15^\circ,\ j_\varphi < 0.15^\circ;$$

$$S_1 : \Delta\varphi_{hg} < 0.15^\circ,\ j_\varphi \geq 0.15^\circ;$$

$$S_2 : \Delta\varphi_{hg} \geq 0.15^\circ,\ j_\varphi < 0.15^\circ;$$

$$S_3 : \Delta\varphi_{hg} \geq 0.15^\circ,\ j_\varphi \geq 0.15^\circ.$$

Thus $\Omega = \{S_0\}$ is the state space that the harmonic gear reducer can work, and $\bar{\Omega} = \{S_1, S_2, S_3\}$ is the state space that the harmonic gear reducer cannot work. Considering the effect of wear, when the operation time is t =8000 hours, the reliability level of the harmonic reducer can be evaluated using the proposed performance reliability simulation method.

Since the manufacturing installation dimensions of flex splines, circular splines, wave generator and bearings include both randomness and fuzziness, the error factors that affect the transmission error and the backlash can be processed as random and fuzzy variables. The random variables in the manufacturing installation error can be regarded as normal distribution and the membership function of fuzzy variables can be determined according to the actual experience and opinions of experts. The distribution parameters of random and fuzzy variables in Eq.(31) and Eq.(32) are presented in Table 1 and 2.

The variables $N, \Delta F_1, \Delta F_2, \Delta\rho_n, \Delta f_a, \Delta F_r$ are defined as fuzzy input variables with the membership functions as follows:

Thus, fuzziness and randomness appear in the structural system simultaneously. An algorithm based on chance theory is designed. The steps in the proposed algorithm are as follows:

Algorithm. Chance theory based performance reliability simulation

Step1: The total simulation number of the system is *N*.

Step2: Set the characteristic parameters of fuzzy and random variables by mechanism designer, as shown in Table 2.

Step3: Refer to the method in Section 4.1 and generate random numbers based on the characteristic parameters as samples of fuzzy and random variables.

Step4: Take the samples into the calculation formulas of error $\Delta\varphi_{hg}$ and j_φ.

Step5: Compare the calculation result with the fault criterion and determine whether the error satisfies the criteria or not. The system is considered as success when criteria is satisfied, otherwise it is failure.

Step6: Repeat steps (3) to (6) *N*-times.

Step7: Use the simulation method of statistical analysis model in Section 4.1 to estimate the number of success *k* via the samples and obtain the reliability by *k*/*N*.

Step8: Estimate the chance theory based reliability using the method of section 4.2.

Using the models and the algorithm in Section 4, the reliability simulation test for a harmonic gear reducer is performed. Firstly, the actual simulation is conducted to decide the critical index and the failure criterion of the system. Then the integrated reliability of the failure and the accuracy error factors are simulated. Using the statistical method, the integrated factors that affect the accuracy of the harmonic gear reducer are considered. After *N*=10000 cycles in simulation, the number of success is 8627. Besides, the chance theory based reliability is obtained to be 0.8491 using Eq.(25). That is, after

Table 1. Random variables

Random variables(unit)	Meaning	Distribution form	Mean value	Coefficient of variation
$E_{rb}(mm)$	The radial motion of the working axis	Normal	0.01	0.050
m	Gear modulus	Normal	0.3	0.033
$\Delta\rho_1\ \Delta\rho_2\ \Delta\rho_3\ \Delta\rho_4\ \Delta\rho_5\ \Delta\rho_6(mm)$	Radial error of the wave generator	Normal	0.01	0.100
E_M	The M value deviation	Normal	0.038	0.132
$u_r(mm)$	Radial clearance of flexible bearings	Normal	0.003	0.0333
$\mu(mm)$	Wear rate	Normal	1.140×10^{-6}	0.136

Table 2. Fuzzy variables

Parameters(unit)	Meaning	Value	Type
N	The number of teeth actually engaged	0.4/0.5/0.6	Fuzzy triangular variable
$\Delta F_i(mm)$	Integrated error of circular spline	0.005/0.075/0.01	Fuzzy triangular variable
$\Delta F_2(mm)$	Integrated error of flex spline	0.005/0.075/0.01	Fuzzy triangular variable
$\Delta\rho_n(mm)$	Integrated radial error of the wave generator	0.01/0.02/0.03/0.04	Fuzzy trapezoidal variable
$\Delta f_a(mm)$	Radial error of the long axis	0.014/0.021/0.028	Fuzzy triangular variable
$\Delta F_r(mm)$	Coaxial error of gear ring	0.009/0.015/0.021	Fuzzy triangular variable

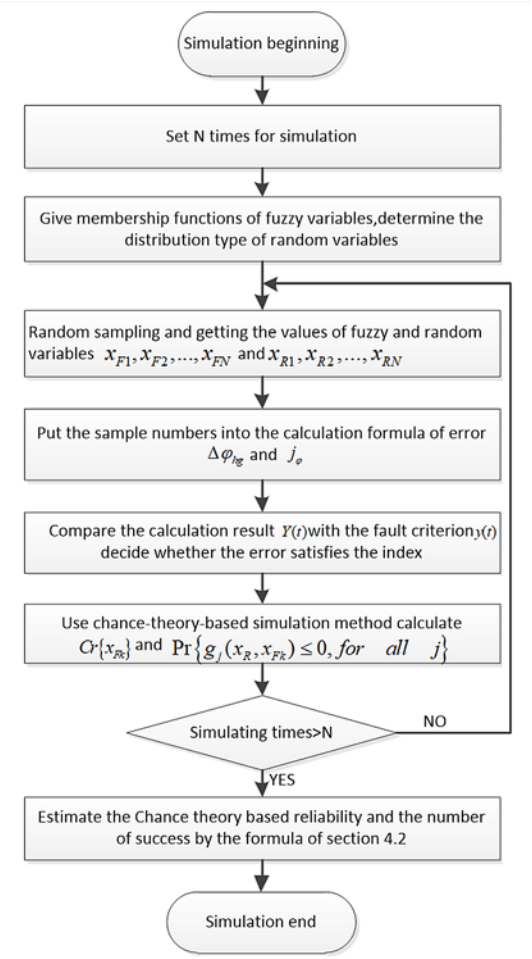


Fig. 5. The flowchart of chance theory based performance reliability simulation

working for 8000 hours, the reliability of harmonic gear reducer drops to 0.8491.

Furthermore, depending on the working time t, and using the above simulation method, a set of corresponding results about the number of success k and the chance theory based reliability can be obtained. Accordingly, the trend of the reliability of the harmonic reducer with time t can be observed. The quadratic polynomial fitting method is used to fit the results. As shown in Fig.6, curve 1 is the result of quadratic polynomial fitting of the statistical analysis, and the curve 2 is the fitting result of the chance based reliability.

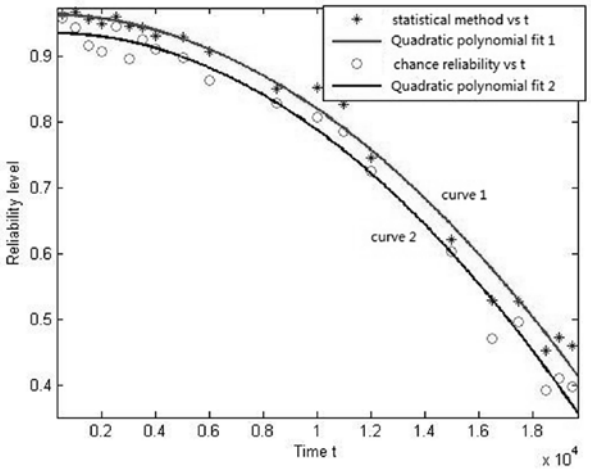


Fig. 6. Time-dependent reliability of the harmonic gear reducer

From the results shown in Table 3 and Fig.6, the chance theory based reliability is 0.9728 at time t=0, which is mainly caused by the manufacturing and installation errors of the flex splines, circular splines, and wave generators of harmonic gear reducers. However, as the working time increases, the wear between the inner wall of the flex spline and the outer wall of the flexible bearing increases with respect to the moving interface, and the reliability decreases gradually. The main reason is that the wear amount gradually accumulates and the wear standard deviation becomes larger over time, which increases the uncertainty of the result. Hence, the influence of the uncertainty on the reliability of the harmonic gear reducer becomes larger. Thus, the results obtained by the chance theory based performance model are consistent with the changes in the reliability of the actual situation affected by the uncertainty factors.

As shown in Fig.6, the trend of the two fitting curves is basically the same. It can be seen from the figure that the reliability obtained by the proposed chance theory based hybrid performance model is smaller than that by the statistical method. Thus, using the statistical method in Section 4.1 to determine the reliability of the structure with hybrid variables will negatively affect the accuracy and the belief degree of the results. That is, the reliability obtained by the proposed method is more conservative and can ensure the structure security to a larger extent. The result of the proposed method presents more realistic estimation of the structure reliability compared to the classic methods.

Table 3. Reliability at different values of time

Time t	Success number (/10000)	$Ch_{reliability}$	Time t	Success number (/10000)	$Ch_{reliability}$
0	0.9860	0.9728	6000	0.9059	0.8638
500	0.9635	0.9569	8500	0.8507	0.8283
1000	0.9666	0.9426	10000	0.8524	0.8058
1500	0.9546	0.9263	11000	0.8261	0.7848
2000	0.9486	0.9358	12000	0.7448	0.7255
2500	0.9588	0.9453	15000	0.6221	0.6037
3000	0.9441	0.8958	16500	0.5296	0.4714
3500	0.9426	0.9239	17500	0.5273	0.4967
4000	0.9307	0.9105	18500	0.4530	0.3930
5000	0.9276	0.8966	19000	0.4604	0.3991

6. Conclusions

In this paper, the chance theory and the multi-state performance reliability model are used to deal with the epistemic uncertainty, especially with the subjective randomness and fuzziness mixed in the mechanism reliability analysis. A chance theory based quantification method integrating the design for performance reliability is proposed. A performance reliability model is established based on fuzzy and random factors related to the structure failure. The hybrid variables and the performance state function are used to represent the subjective randomness and fuzziness. The chance measure is adopted to quantify the reliability of the structures. In the framework of the chance theory, a definition of the chance theory based performance reliability is provided and a formulation is proposed to analyze the reliability in a structural system from a viewpoint of the chance measure. With the proposed formulations, not only the duality of random events can be described, but also the subadditivity of fuzzy events can be explained. A harmonic gear reducer considering wear failure is provided as a case in order to demonstrate and validate the correctness and effectiveness of the proposed algorithm. The results obtained from the case experiment have demonstrated that the trend of the chance theory based reliability is basically consistent with that of the statistical analysis results. Moreover, the comparative results have shown that the results of the proposed chance theory based simulation method are more conservative and this method can reflect the influence of the hybrid variables on the structural reliability more accurately than the classical methods. Thus, as a new method to deal with fuzzy and random hybrid uncertainties, the proposed chance theory based reliability model can effectively solve the existing problems in engineering practice.

Acknowledgement:
This study was supported by the National Natural Science Foundation of China (No. 51675026, 71671009), and the National Basic Research Program of China (No. 2013CB733002).

References

1. Breitung K. 40 years FORM: Some new aspects? Probabilistic Engineering Mechanics 2015; 42:71-77, https://doi.org/10.1016/j.probengmech.2015.09.012
2. Chen Y, Wen M, Kang R. Model and Theory for Fuzzy-Based Performance-Reliability. International Information Institute (Tokyo). Information 2013; 02; 16(2):951-959.
3. Cai K Y. Introduction to Fuzzy Reliability. Springer US,1996, https://doi.org/10.1007/978-1-4613-1403-5.
4. Cremona C, Gao Y. The possibilistic reliability theory: theoretical aspects and applications. Structural Safety 1997; 19(2):173-201, https://doi.org/10.1016/S0167-4730(97)00093-3.
5. Chakraborty S, Sam P C. Probabilistic safety analysis of structures under hybrid uncertainty. International journal for numerical methods in engineering 2007; 70(4): 405-422, https://doi.org/10.1002/nme.1883.
6. Chen Y, Kang R, Sun Y. Integrating design for performance and reliability. Reliability and Maintainability Symposium 2006. RAMS'06. Annual. IEEE 2006: 343-348.
7. Eldred M S, Swiler L P, Tang G. Mixed aleatory-epistemic uncertainty quantification with stochastic expansions and optimization-based interval estimation. Reliability Engineering & System Safety 2011; 96(9): 1092-1113, https://doi.org/10.1016/j.ress.2010.11.010.
8. Gnedenko B, Ushakov I A. Probabilistic reliability engineering. John Wiley & Sons 1995, https://doi.org/10.1002/9780470172421.
9. Huang H Z. Structural reliability analysis using fuzzy sets theory. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14: 284-294.
10. Holický M. Fuzzy probabilistic models in structural reliability[J]. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2006; 2:11-13.
11. Kang R, Zhang QY, Zeng ZG, Zio E, Li XY. Measuring reliability under epistemic uncertainty: Review on non-probabilistic reliability metrics. Chinese Journal of Aeronautics 2016; 29(3):571–9, https://doi.org/10.1016/j.cja.2016.04.004.
12. Kaufmann A. Introduction to the theory of fuzzy subsets. Academic Pr 1975.
13. Kapur K C. Multi-state reliability: models and applications. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2006; 2: 8-10.
14. Li G, Lu Z, Li L, et al. Aleatory and epistemic uncertainties analysis based on non-probabilistic reliability and its kriging solution. Applied Mathematical Modeling 2016; 40(9-10): 5703-5716, https://doi.org/10.1016/j.apm.2016.01.017.
15. Liu B, Liu Y K. Expected value of fuzzy variable and fuzzy expected value models. IEEE Transactions on Fuzzy Systems 2002; 10(4):445-450, https://doi.org/10.1109/TFUZZ.2002.800692.
16. Liu B. Uncertainty theory: An introduction to its axiomatic foundations. Berlin: Springer 2004, https://doi.org/10.1007/978-3-540-39987-2.
17. Li X, Liu B. A sufficient and necessary condition for credibility measures. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 2006; 14(05): 527-535, https://doi.org/10.1142/S0218488506004175.
18. Liu B. A survey of credibility theory. Fuzzy optimization and decision making 2006; 5(4): 387-408. https://doi.org/10.1007/s10700-006-0016-x
19. Liu B. Uncertainty Theory (2nd ed.). Berlin: Springer 2007, https://doi.org/10.1007/978-3-540-73165-8_5.
20. Li L, Lu Z. Interval optimization based line sampling method for fuzzy and random reliability analysis. Applied Mathematical Modeling 2014; 38(13): 3124-3135, https://doi.org/10.1016/j.apm.2013.11.027.
21. Li X, Liu B. Chance measure for hybrid events with fuzziness and randomness. Soft Computing 2009; 13(2): 105, https://doi.org/10.1007/s00500-008-0308-x.
22. Liu B. Uncertainty theory: A branch of mathematics for modeling human uncertainty. Berlin: Springer 2010, https://doi.org/10.1007/978-3-642-13959-8.
23. Liu B. Theory and practice of uncertain programming (2nd ed.). Berlin: Springer 2009, https://doi.org/10.1007/978-3-540-89484-1.
24. Liu B. A survey of entropy of fuzzy variables. Journal of Uncertain Systems 2007; 1(1): 4-13.
25. Liu B. Random fuzzy dependent-chance programming and its hybrid intelligent algorithm. Information sciences 2002; 141(3): 259-271, https://doi.org/10.1016/S0020-0255(02)00176-7.

26. Liu B. Uncertainty Theory (4th ed.). Berlin: Springer 2015, https://doi.org/10.1007/978-3-662-44354-5.
27. Li J Y, Wang J X, Zhou G W, et al. Accelerated life testing of harmonic driver in space lubrication. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology 2015; 229(12): 1491-1502, https://doi.org/10.1177/1350650115586032.
28. Liu Y. Uncertain random variables: a mixture of uncertainty and randomness. Soft Computing 2013; 17(4): 625-634, https://doi.org/10.1007/s00500-012-0935-0;
29. Marano G C, Quaranta G. A new possibilistic reliability index definition. Acta Mechanica 2010; 210 (3-4): 291-303, https://doi.org/10.1007/s00707-009-0194-z.
30. Muscolino G, Santoro R, Sofi A. Reliability analysis of structures with interval uncertainties under stationary stochastic excitations. Computer Methods in Applied Mechanics and Engineering 2016; 300: 47-69, https://doi.org/10.1016/j.cma.2015.10.023
31. Murchland J D. Fundamental concepts and relations for reliability analysis of multi-state systems. Reliability and fault tree analysis 1975.
32. Utkin L V, Gurov S V. A general formal approach for fuzzy reliability analysis in the possibility context. Fuzzy Sets & Systems 1996; 83(2):203-213, https://doi.org/10.1016/0165-0114(95)00391-6.
33. Wang P, Zhang J, Zhai H, et al. A new structural reliability index based on uncertainty theory. Chinese Journal of Aeronautics 2017, https://doi.org/10.1016/j.cja.2017.04.008.
34. Wang Z, Huang H Z, Li Y, et al. An approach to system reliability analysis with fuzzy random variables. Mechanism & Machine Theory 2012; 52(52):35-46, https://doi.org/10.1016/j.mechmachtheory.2012.01.007.
35. Wen M, Kang R. Reliability analysis in uncertain random system. Fuzzy Optimization and Decision Making 2016; 15(4): 491-506, https://doi.org/10.1007/s10700-016-9235-y.
36. Zadeh L A. Fuzzy sets, information and control. Information & Control 1965; 8(3): 338-353, https://doi.org/10.1016/S0019-9958(65)90241-X.
37. Zadeh L A. Fuzzy sets as a basis for a theory of possibility. Fuzzy sets and systems 1978; 1(1): 3-28, https://doi.org/10.1016/0165-0114(78)90029-5.

Lei ZHANG
Jianguo ZHANG
Hao ZHAI
Shuang ZHOU
Science and Technology on Reliability and Environmental Engineering Laboratory
Beihang University
Xueyuan Road No.37, Haidian District, Beijing 100083, China
School of Reliability and Systems Engineering
Beihang University
Xueyuan Road No.37, Haidian District, Beijing 100083, China

E-mails: zhanglei16@buaa.edu.cn, zjg@buaa.edu.cn, zhaihao2010@126.com, zsdyx88@163.com

Article citation info:
WOLAK A, ZAJĄC G, KUMBÁR V. Evaluation of engine oil foaming tendency under urban driving conditions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 229–235, http://dx.doi.org/10.17531/ein.2018.2.07.

Artur WOLAK
Grzegorz ZAJĄC
Vojtěch KUMBÁR

EVALUATION OF ENGINE OIL FOAMING TENDENCY UNDER URBAN DRIVING CONDITIONS

OCENA ZMIAN CHARAKTERYSTYK PIENIENIA OLEJÓW SILNIKOWYCH UŻYTKOWANYCH W WARUNKACH JAZDY MIEJSKIEJ

The purpose of the article was to analyze the foaming tendency of engine oils used under excessive operating conditions. To achieve this end, foaming characteristics were determined for 23 oil samples in three measurement sequences. Foaming tendency was measured using the ASTM D 892 standard method, which consists in assessing foaming tendency of the liquid and foam stability. The cars used in the tests were uniform in terms of brand, type and operating conditions. The relationship between the mileage of the cars tested and the volume as well as stability of foam in used engine oils were presented using scatter plots with regression lines, correlation coefficient and 95% confidence interval. Based on the obtained results it was found that foaming tendency for new oils is characterized by high variability. The strongest foaming tendency at 24°C and 93°C (Sequence I and II) was observed for two out of five oil groups. Statistically significant differences were found between mileage and foaming tendency/foam stability for individual oils tested.

***Keywords**: degradation, engine oil, foam tendency, foam stability, oil condition monitoring.*

Celem opracowania była analiza zmian charakterystyk pienienia olejów silnikowych użytkowanych w warunkach jazdy miejskiej. W związku z realizacją sformułowanego celu oznaczono skłonność do pienienia oraz trwałość piany dla 23 próbek oleju, według normy ASTM D 892. Próbki oleju pochodziły z samochodów stanowiących jednolitą flotę pod względem: marki, typu oraz warunków pracy silnika. Zależności przebiegu eksploatacyjnego badanych samochodów, z ilością i trwałością piany, w przepracowanych olejach silnikowych sprawdzono poprzez zastosowanie wykresów rozrzutu z linią regresji, współczynnikiem korelacji oraz 95%-owym przedziałem ufności. Na podstawie uzyskanych wyników stwierdzono, że skłonność do tworzenia piany dla olejów świeżych charakteryzuje się dużym zróżnicowaniem. Największą skłonność do pienienia w temp. 24°C i 93°C (Sekwencja I i II) zaobserwowano dla dwóch (z pięciu) grup olejowych. Potwierdzono istotne statystycznie różnice pomiędzy przebiegiem badanych pojazdów a poziomami poszczególnych charakterystyk pienienia.

***Słowa kluczowe**: degradacja, olej silnikowy, skłonność do pienienia, trwałość piany, monitorowanie jakości oleju.*

1. Introduction

Belonging to the group of operating fluids, engine oils are generally known to exhibit foaming tendencies. This particular property of oils may cause a number of unfavorable effects in actual use, such as: too low amount of lubricant applied to the brake assembly, increased compressibility of operating fluids (resulting from the presence of air bubbles), heat dissipation difficulties or accelerated oxidation. Moreover, foaming can result in poor system performance and can cause serious mechanical damage: disturbances in the lubrication and cooling of the cooperating elements, the decrease in the pump capacity or the discharge of the operating fluid from the tank of the machine which, in turn, can lead to an increase in maintenance costs and pollution of the environment. That is why foam resistance is considered one of the most important properties of lubricants [4, 6-8].

The foaming tendencies of engine oils represent a serious problem, especially when they are stronger than what might be deemed permissible under given operating conditions. Resistance of the operating fluids to foaming depends on many factors, including the chemical and physical properties of these fluids. The influence of viscosity, density and surface tension on this phenomenon is very significant. It is worth noting that unfavorable operating conditions and poorly structured oil circuits are particularly conducive to foam formation.

Engine oil which should create only limited amounts of foam is subjected to foaming resistance tests. Foaming characteristics are measured by the ASTM D 892 standard method, which consists in evaluating the foaming tendency of the liquid and foam stability. The lower amount of foam, the better resistance to foaming. That is why, most lubricants contain antifoam additive to break up foams [1, 10].

Attempts to determine the limit levels for engine oils, found in both subject literature and practice [5, 9-13], all come down to assessing the permissible changes of the following [2, 12]:

- one selected physicochemical property of oil,
- one selected engine parameter,
- a parameter considered to be representative of the aging process,
- a set of parameters characterizing oil quality,
- a concentration of additives,
- the value of the synthetic parameter,
- similarities in the condition of oils.

Individual oil condition indicators (viscosity, acid number, base number, etc.) are widely used; nevertheless, it is equally important to consider a more general indicator which is the oil's tendency to foam.

The purpose of this paper is to analyze the foaming tendencies of engine oils used under excessive operating conditions. To achieve this end, the foaming characteristics (foaming tendency and foam stability) were determined for 23 oil samples in three measurement sequences, as described in the subject guidelines.

2. Materials and methods

The lubricating oils selected for testing have all complied with the specifications of Mitsubishi Motors Corporation (the manufacturer of the engines used in the study). The research material comprised of engine oil samples coded as CE, MS, ME, PS and PE. The quality of engine oils used in cars of the same type, operated under the same conditions and over the same period of time, was thoroughly assessed. The cars used in the tests were uniform in terms of brand, type and operating conditions and belonged to the Driver Training Center fleet. In total, there were 23 vehicles with spark ignition engines (cylinder capacity of 1,332 cm^3) and operating on fuel coming from the same producer. Detailed specifications of the engine are presented in Table 1. All of the cars were generally used for conducting driving license tests. Three out of five oils (MS, ME and PE) were applied to five cars each, whereas two out of five (CE and PS) were applied to four cars each. At the beginning of the investigation an unused/new oil sample was tested. Then, the samples were collected and examined after one year (when the oil was changed).

Table 1. Detailed specifications of the engine.

Model code	4A90
Type	DOHC MIVEC
Cylinder volume [cm^3]	1,332
Maximum power [kW@rpm]	80@6,000
Maximum torque [Nm@rpm]	145@4,000
Fuel supply	MPI
Capacity of the lubrication system	4 l.
Required engine oil specifications	Synthetic 5W-30

The engine oil specifications (grade SAE 5W-30) are presented in Table 2 together with the API and ACEA standards.

Table 2. Quality and viscosity grades of the engine oils selected for testing.

Classification	Oil				
	CE	ME	MS	PE	PS
SAE	5W-30	5W-30	5W-30	5W-30	5W-30
ACEA	A3/B3-10, C3-10	C2/C3-10	C3-10	C2-10	A3/B3-10
API	SH	SM/SN	SM/SL	-	SM

Details on the number of kilometers traveled by each car selected for testing can be found in Table 3. The sample code includes the oil code and the car code. When analyzing the results, account should be taken of "harsh" operating conditions of the vehicles, i.e. operating at varying engine load, extended idling, driving in street traffic, short distance driving causing under-heating of the engine, and frequent start-ups at low engine temperatures. It should be noted that the tests as well as the sampling procedure were planned (in agreement with the fleet owners) in such a way as to avoid topping up the oil between the test periods. Ultimately, all selected cars were used for examination purposes, but after the completion of the study, it turned out that two of the cars from the selected fleet were actually used for other purposes. The one with the engine oil coded PS 018193 was used by the employees of the Driver Training Center as a company car. The second car, to which the engine oil CE 017977 was applied, was used for training purposes. Both cars practically never served the examination purposes. Such information is a great contribution to this work, as it also demonstrates the impact of the operating conditions on the properties of engine oils.

Table 3. The number of kilometers traveled by the cars tested.

Sample code	Mileage – starting point [km]	The number of kilometers traveled [km]
CE 17760	12858	13220
CE 17943	14571	13907
CE 17977*	**6033**	**9975**
CE 18716	9477	12169
ME 17764	15484	14501
ME 17810	7840	8749
ME 17973	8015	9029
ME 18345	9753	9234
ME 18760	15307	14573
MS 18011	13486	9005
MS 18128	14571	10803
MS 18361	14988	14166
MS 18793**	**6223**	**6100**
MS 18817	14666	13877
PE 17939	12370	13066
PE 18024	11694	11748
PE 18207	12209	13223
PE 18591	15409	15188
PE18689	13290	15531
PS 18149	13556	14384
PS 18193*	**14457**	**18211**
PS 18784	11948	12317
PS 18799	14154	13523

** – used for different purposes/operated under different conditions,*
*** – significantly lower mileage*

The standard method generally used is to measure the volume of foam generated by blowing air through the sample of the test oil in the period of 5 minutes (foaming tendency), followed by a measurement of the foam volume after allowing it to stand for 10 minutes or by a measurement of the total time until the foam completely disappears. The measured volume of foam after this time is called the "foam stability".

The lab set for measuring the resistance of oil to foaming (Fig. 1) consists of:

- a measuring cylinder made of glass resistant to heating, of volume 1000 cm^3, and the scale interval of 10 cm^3,
- a rubber stopper (fitted to the cylinder) with two holes: the central one to insert the sparger and lateral for the placement of the air discharge tube,

Fig. 1. The workstation where the foaming tendency of oils was measured; Department of Performance Testing, Oil and Gas Institute in Kraków

- a sparger, consisting of a brass air supply pipe and a spherical aluminum oxide membrane filter, cm^3,
- a lab bath, with thermo-regulators, filled with water,
- a flow meter to measure air volume flow,
- a thermometer with a measuring range of 0-100°C,
- a thermo-regulator - an electric heater with automatic temperature control.

The measurements were made in three sequences:
- sequence I – measurement at the temperature of 24°C,
- sequence II – measurement at the temperature of 93°C (a new sample),
- sequence III – measurement at the temperature of 24°C performed after the measurement at the temperature of 93°C.

The requirements regarding foaming tendency are given in the following way:
- **sequence I – no more than 50/10 [foam volume, cm^3]**, (interpretation: at the temperature of 24°C the foaming tendency is of 50 cm^3, whereas the foam stability is of 10 cm^3),
- **sequence II – no more than 100/10 [foam volume, cm^3],** (interpretation: at the temperature of 93°C the foaming tendency is of 100 cm^3, whereas the foam stability is of 10 cm^3),
- **sequence III – no more than 50/10 [foam volume, cm^3],** (interpretation: at the temperature of 24°C the foaming tendency is of 50 cm^3, whereas the foam stability is of 10 cm^3) – measurement at the temperature of 24°C performed for the sample after the measurement at the temperature of 93°C.

The first figure represents the requirements for the foaming tendency and the second for foam stability.

The results obtained were statistically analyzed using the STATISTICA software (StatSoft Inc., USA). To investigate the relationship between the mileage of the cars tested and the volume as well as stability of foam in used engine oils, scatter plots with regression lines, correlation coefficient and 95% confidence interval were used. Shapiro-Wilk test was applied to assess whether data are normally distributed.

3. Results and discussion

Detailed results for each sample tested are shown in Table 4. The results for sequence I, II, and III measurements are summarized in each column. In sequence I, at the temperature of 24°C the strongest foaming tendency was observed in three samples from the MS oil group (MS 018011, MS 018361, MS 018128 with the following values – 680, 670, and 650 cm^3, respectively) and in two oil samples from the ME group (ME 018345, ME 017973 with the values of 640 and 630 cm^3). The same samples showed the highest foam stability measured at the temperature of 24°C. It was on the level of 490 cm^3 (an average for the three MS oil group samples) and 345 cm^3 (an average for the two ME oil group samples). In sequence II, at the temperature of 93°C, the oil samples from the MS group (MS 018011, MS 018361, MS 018128 – an average: 570 cm^3), continued to exhibit the strongest foaming tendency. Other samples with relatively high foaming tendency were those from the ME oil group (ME018345 and ME 017973 – an average: 395 cm^3) and from the PS oil group (PS 018149, PS 018784 and PS 018799 – an average: 370 cm^3). It is worth noting that in spite of various amounts of foam in samples at 93°C, in all analyzed cases, the foam stability 10 minutes after discontinuing the air blow was at zero level. In sequence III, the foaming tendency and foam stability were measured again at the temperature of 24°C; the difference was that it was a measurement of the samples that had been tested at 93°C before. The results are consistent with those obtained in sequence I. In the PS oil group, it is worth analyzing further the oil sample PS 018193 and in the CE group – the oil CE 017977, which despite the large number of kilometers traveled have retained the lowest levels of foam volume and stability. This clearly proves that unfavorable operating conditions increase the foaming tendency of oils.

When analyzing the results of the foaming tendency tests (Table 4) for new oils (before application to the engine) it can be observed that only one oil type (CE) did not exhibit foaming tendency in any of the three sequences, the MS oil showed some foaming tendency in sequence II, whereas the remaining oils exhibited foaming tendencies in all three sequences. The strongest foaming tendency was observed for the PE oil (SI-150 cm^3, SII-50 cm^3 and SIII-90 cm^3), yet after the service life ended, the average foaming tendency of this oil was similar to that of the CE oil group and significantly lower than in the other three groups (which actually had much lower foaming tendency and foam stability for new oil than the PE oil).

Figures 2 and 3 illustrate the results of the foaming tendency measurements. The symbols (in boxes) indicate the average values of the individual parameters, while the upward and downward whiskers depict the distribution of individual measurements around the mean (standard deviation).

The strongest foaming tendency in sequence I was observed for the ME oil group samples ($\overline{x}$ = 600 cm^3, s = 33 cm^3), in sequence II for the MS oils ($\overline{x}$ = 458 cm^3, s = 149 cm^3) and in sequence III for the ME oil group again ($\overline{x}$ = 600 cm^3, s = 6 cm^3). The lowest mean values were found in all three sequences for the CE oil group samples – in SI ($\overline{x}$ = 295 cm^3, s = 166 cm^3), in SII ($\overline{x}$ = 255 cm^3, s = 11 cm^3) and in SIII ($\overline{x}$ = 403 cm^3, s = 52 cm^3). High values of standard deviation in the CE, MS and PS oil groups are caused by different operating conditions under which the oils PS 018193 and CE 017977 were used and a relatively low mileage of the car with the oil MS 018793. Excluding these three

Table 4. The foaming tendency and foam stability of the oil samples tested

Sample code	Sequence 1 (SI)	Sequence 2 (SII)	Sequence 3 (SIII)
CE0K	0/0	0/0	0/0
CE 017760	420/170	250/0	480/200
CE 017943	360/180	270/0	360/150
CE 017977*	10/0	240/0	350/0
CE 018716	390/170	260/0	420/180
ME0K	10/0	60/0	20/0
ME 017764	600/430	340/0	600/400
ME 017810	550/300	310/0	590/290
ME 017973	630/340	390/0	600/310
ME 018345	640/350	400/0	610/330
ME 018760	580/320	330/0	600/310
MS0K	0/0	20/0	0/0
MS 018011	680/460	580/0	620/450
MS 018128	650/490	570/0	620/460
MS 018361	670/520	550/0	610/510
MS 018793**	30/0	190/0	340/30
MS 018817	520/330	400/0	570/470
PE0K	150/0	50/0	90/0
PE 017939	490/70	170/0	460/60
PE 018207	340/50	160/0	360/40
PE 018024	480/80	190/0	440/50
PE 018591	490/100	240/0	500/60
PE 018689	510/110	230/0	520/70
PS0K	50/0	70/0	30/0
PS 018149	550/180	380/0	480/170
PS 018193*	20/0	270/0	340/20
PS 018784	510/90	370/0	500/170
PS 018799	540/160	360/0	470/160

** – used for different purposes/ operated under different conditions*

*** – significantly lower mileage*

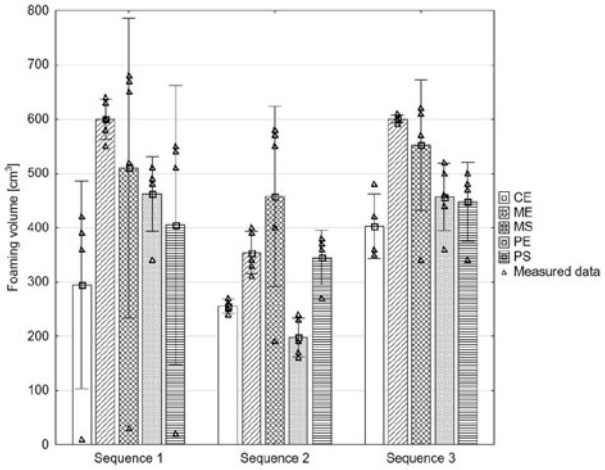


Fig. 2. The foaming tendency of engine oils measured after their service life ended (with mean values and standard deviations)

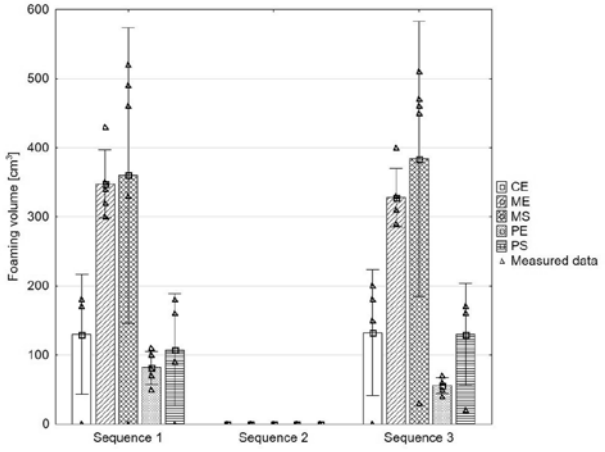


Fig. 3. The foam stability of engine oils measured after their service life ended (with mean values and standard deviations)

samples from the mean and standard deviation, the following results are obtained: in SI for the CE oils ($\overline{x}$ = 390 cm^3, s = 24 cm^3), for the MS oil group ($\overline{x}$ = 630 cm^3, s = 64 cm^3) and for the PS oil group ($\overline{x}$ = 533 cm^3, s = 17 cm^3). In SII for the CE oils ($\overline{x}$ = 260 cm^3, s = 8 cm^3), for the MS oils ($\overline{x}$ = 525 cm^3, s = 73 cm^3) and for the PS oil group ($\overline{x}$ = 370 cm^3, s = 8 cm^3). In SIII for the CE oil samples ($\overline{x}$ = 420 cm^3, s = 49 cm^3), for the MS oil group ($\overline{x}$ = 605 cm^3, s = 21cm^3) and for the PS oil group ($\overline{x}$ = 483 cm^3, s = 12 cm^3).

The highest mean value of foam stability in sequence I and in sequence II was observed for the MS group of oils. It was ($\overline{x}$ = 360 cm^3, s = 191 cm^3) and ($\overline{x}$ = 384 cm^3, s =178 cm^3), respectively. Slightly lower values (ca. 12 cm^3 lower in sequence I and ca. 56 cm^3 lower in sequence III) were observed for the MS oil group. The lowest mean values were noted for the PE oil group – in SI ($\overline{x}$ = 82 cm^3, s = 21 cm^3) and in SIII ($\overline{x}$ = 56 cm^3, s = 10 cm^3). Similar tendencies were observed for CE and PS oil groups (the mean foam stability in both sequences for these groups is 125 cm^3). Also, with regard to foam stability, particularly high standard deviations are noticeable in CE, MS and PS oil groups due to different operating conditions of some of the oils (PS 018193 and CE 017977) and relatively low mileage of the car with the MS 018793 oil. Excluding these three samples from the mean and standard deviation calculation, the following results are obtained: in SI for the CE oil group ($\overline{x}$ = 173 cm^3, s = 5 cm^3), for the MS oil group ($\overline{x}$ = 450 cm^3, s = 72 cm^3) and for the PS oil group ($\overline{x}$ = 143 cm^3, s = 39 cm^3). In SIII for the CE oil group ($\overline{x}$ = 177 cm^3, s = 21 cm^3), for the MS oil group ($\overline{x}$ = 473 cm^3, s = 23 cm^3) and for the PS oil group ($\overline{x}$ = 167 cm^3, s = 5 cm^3).

The test results for engine oils operated under similar driving conditions show that the samples from the PE oil group exhibit lower foaming tendency at the temperature of 93°C, or are characterized by a greater ability to eliminate foam. On the other hand, the MS oil group samples are also worth mentioning. In their case, when analyzing the results in sequence II, a large number of emerging air bubbles can significantly reduce not only the lubrication capacity due to the presence of air but also the efficiency of the pumping systems. It should also be noted that the impact of car use and operating conditions on the foaming tendency of oils is significant. As it has been previously mentioned, two of the cars in the test were used in urban driving conditions, whereas the other ones were operated under "harsh" conditions (short distance driving, under-heating of the engine etc.)

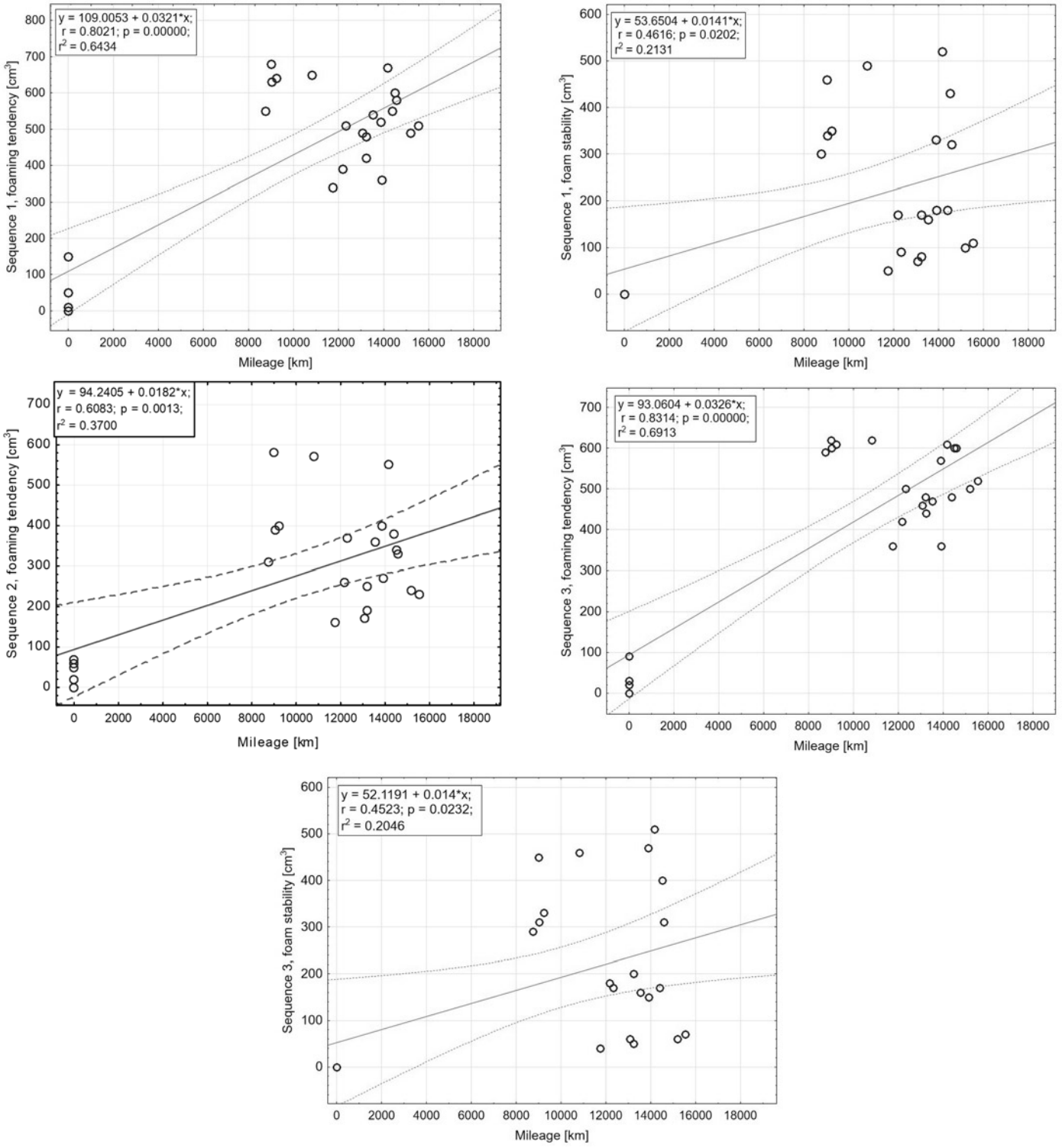


Fig 4. Scatter plots with regression lines, correlation coefficients, and 95% confidence intervals

because they were generally used for conducting driving license tests. The study results indicate that for the two vehicles operated differently, both the foaming tendency and foam stability significantly differ from the properties of the other oils tested. It is particularly noticeable in the foam stability for sequence I and III, which remains stable on the zero level, whereas it exceeds the value of 100 cm^3 in all other oils. This may be related to the fact that more severe operating conditions accelerate the degradation of oil. According to Duncanson [3], water, solid impurities and oxidation products can contribute to the creation of foam in lubricating oils.

At the same time, it can be observed that also the service-life of oil may affect its foaming tendency – the higher the mileage, the stronger foaming tendency and foam stability. This trend is visible for groups of vehicles operated under similar driving conditions. With prolonged engine oil use, the products that result from the oxidation of oils become the dominant influence on the foaming tendency. After some period of oil use, a large amount of oxidation products is formed which settle on the particles and convert them into a colloidal suspension. Such colloidal suspension has solid particles surrounded

by compounds such as: resins, asphaltenes and macromolecular acidic substances, which become foam stabilizers.

Assuming similar operating conditions (and excluding the following oils: PS 018193 and CE 017977), such significant differences in the foaming tendency as observed for individual oil groups may be explained by differences in the amount of antifoam additives or by the type of additives used by manufacturers.

According to expert tests [9], the critical value for the foam volume at 93°C is 200 cm^3, whereas for foam stability it is max. 30 cm^3. Upon analyzing the results obtained for the limit values, it should be noted that while the foam stability remains at zero level, in the case of the amount of foam accumulated as a result of blowing air through the test oil sample in the period of 5 minutes, there are significant differences observed. The PE oils, for which the average value remains at the permissible limit, are definitely the best performers. The other oils tested exhibit very strong foaming tendencies with the highest values achieved by the ME and MS oils.

In order to check whether there is any statistical relationship between the mileage and the foaming tendency/foam stability, scatter plots with regression line and correlation coefficient were used (Fig. 4). It should be noted that three samples (PS 018193, CE 017977, MS 018793) have been left out in statistical analyzes due to their not meeting the prerequisite of being operated under similar conditions.

When analyzing the received significance levels (0.0000; 0.0202; 0.0013; 0.0232), it was found that all of them are below the limit value (0.05), so they should be considered statistically significant. It may be therefore concluded that there are also statistically significant differences between the mileages of the cars tested and the individual foaming characteristics (except for SII, foam stability). The obtained correlation coefficients show that in all five analyzed cases positive correlation was obtained. This, in turn, suggests that an increase in the foam volume and stability measured at the temperatures of 24 and 93°C (SI and SII) is related to the vehicle's service life, expressed in the number of kilometers traveled. The highest correlation coefficient was obtained for the foaming tendency at 24/93°C (SIII, r = 0.8314). Using J. Guilford's terminology, it is indeed a high correlation. Scatter plots with regression lines and correlation coefficients also show strong outliers. These are PS 018193 and CE017977 oils that have been used in cars operated under different conditions.

The observed increase in foaming tendency mainly results from gradual depletion of anti-foaming agents in oil. It may also be affected by physicochemical properties of oil. In order to obtain the full picture parameters such as: kinematic viscosity, HTHS, CCS, TAN, oxidation, nitration, water content, wear debris content were also measured. However, the research material obtained is very extensive and the changes in the abovementioned parameters have become the subject of other papers that have already been published [11, 12] or are currently in print.

4. Conclusions

The creation of foam in engine oils negatively affects the entire lubrication system and may increase the wear of engine elements thus contributing to reduced engine life. Engine oils should have high foam resistance and low foam stability. This is achieved by the selection of suitable oil components and anti-foam additives. However, there are several factors that may change the characteristics of oil foaming during operation. The conducted research on a fleet of 23 vehicles operated under similar conditions allowed us to compare the foaming characteristics of oils after their service life ended. Based on the results obtained, it was found that:

- Foaming tendency for new oils is characterized by high variability. Only new oil from the CE group did not exhibit foaming tendencies in any of the sequences. The MS oil exhibited foaming tendency in sequence II, whereas the other oils showed foaming tendencies in all three sequences. The PE oil showed the strongest foaming tendency among new oils tested (SI – 150 cm^3, SII – 50 cm^3 and SIII – 90 cm^3).
- The results of in-service research carried out under similar conditions show that the PE oils are either less likely to exhibit foaming tendency or have a higher foam elimination capacity. An inverse relationship was obtained, compared to new oils.
- The strongest foaming tendency at 24°C (SI) was observed for three samples from the MS oil group (MS 018011, MS 018361, MS 018128 and these were the following values – 680, 670, and 650 cm^3, respectively) and for two samples from the ME oil group (ME 018345 and ME 017973 with the respective values of 640 and 630 cm^3).
- The same samples showed the highest foam stability measured at 24°C (SI) - on the level of 490 cm^3 (an average for three MS oils) and 345 cm^3 (an average for two ME oils).
- The strongest foaming tendency at 93°C (SII) was observed for the MS oil group (MS 018011, MS 018361, MS 018128 – average 570 cm^3), for the ME oil group (ME 018345 and ME 017973 – average 395 cm^3), and for the PS oils (PS 018149, PS 018784 and PS 018799 – average 370 cm^3).
- In all analyzed oil samples, the foam stability for sequence II (10 minutes after the air blow was discontinued) remains at zero level.
- Statistically significant differences were found between the mileage and the levels of individual foaming characteristics ($p < 0.05$).
- Positive correlation coefficients were obtained, thus proving that an increase in the foam volume and stability measured at the temperatures of 24 and 93°C (SI, SII and SIII) is related to the vehicle's service life and expressed in the number of kilometers traveled.

Acknowledgements

This research was supported and financed by project TP 2/2017 "Effect of additives on the rheological behaviour of foodstuffs and raw materials for their production" of Internal Grant Agency FA MENDELU.

All laboratory tests for this study were conducted at the Oil and Gas Institute in Kraków – the National Research Institute. The authors wish to thank the Director of the Oil and Gas Institute: Prof. Maria Ciechanowska, PhD. Eng. and acting as Deputy Director for Oil Technology: MSc Wiesława Urzędowska, for the opportunity to perform the tests.

References

1. Bart J C J, Gucciardi E, Cavallaro S. Lubricant use and disposal. Biolubricants 2013: 755–823, https://doi.org/10.1533/9780857096326.755.
2. Chmielewski Z. Badania oleju silnikowego jako źródła informacji o stanie technicznym silnika spalinowego o ZS. LOGITRANS - VII Konf. Nauk. - Logistyka, Syst. Transp. Bezpieczeństwo w Transporcie 2010.
3. Duncanson M. Effects of physical and chemical properties on foam in lubricating oils. Lubr. Eng. 2003; 59: 9–13.
4. Fowle T I. Aeration in lubricating oils. Tribol. Int. 1981; 14: 151–157, https://doi.org/10.1016/0301-679X(81)90062-1.
5. Kral J, Konecny B, Madac K, Fedorko G, Molnar V. Degradation and chemical change of longlife oils following intensive use in automobile

engines. Measurement 2014; 50: 34–42, https://doi.org/10.1016/j.measurement.2013.12.034.
6. Nemoto S. A study of engine oil aeration. JSAE Rev. 1997; 18: 271–276, https://doi.org/10.1016/S0389-4304(97)00008-8.
7. Prolic T C, Lepusic A. Effect of foaming on the antiwear properties of lubricating oils. Goriva I Maz 2012; 51: 38.
8. Qiao X, Li P, Bai J, Zhuang J, Huang Z. Oil Aeration Measurement on a High-Speed Diesel Engine. SAE International 2014, https://doi.org/10.4271/2014-01-2786.
9. Urzędowska W, Stępień Z. Wybrane zagadnienia dotyczące zmian właściwości silnikowego oleju smarowego w eksploatacji. Nafta-Gaz 2012; 12 (LX):1102–1110.
10. Vasanthan B, Devaradjane G, Shanmugam V. Online condition monitoring of lubricating oil on test bench diesel engine & vehicle. J. Chem. Pharm. Sci. 2015; 4:315–20.
11. Wolak A. TBN Performance Study on a Test Fleet in Real-World Driving Conditions Using Present-Day Engine Oils. Measurement 2018; 114: 322–331, https://doi.org/10.1016/j.measurement.2017.09.044.
12. Wolak A, Zając G. The kinetics of changes in kinematic viscosity of engine oils under similar operating conditions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19(2): 260–267, https://doi.org/10.17531/ein.2017.2.15.
13. Zając G, Szyszlak-Bargłowicz J, Słowik T, Kuranc A, Kamińska A. Designation of chosen heavy metals in used engine oils using the XRF method. Polish J. Environ. Stud. 2015; 24(5): 2277-2283, https://doi.org/10.15244/pjoes/58781.

Artur WOLAK
Department of Industrial Commodity Science
Cracow University of Economics
Sienkiewicza 4, 30-033 Kraków, Poland

Grzegorz ZAJĄC
Department of Power Engineering and Transportation,
Faculty of Production Engineering
University of Life Sciences in Lublin
Głęboka 28, 20-612 Lublin, Poland

Vojtěch KUMBÁR

Department of Technology and Automobile Transport
Mendel University in Brno
Zemědělská 1/1665, 613 00 Brno, Czech Republic

E-mails: artur.wolak@uek.krakow.pl,
grzegorz.zajac@up.lublin.pl, vojtech.kumbar@mendelu.cz

Article citation info:
SONG H, SCHNIEDER E. Modeling of railway system maintenance and availability by means of colored Petri nets. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 236–243, http://dx.doi.org/10.17531/ein.2018.2.08.

Haifeng SONG
Eckehard SCHNIEDER

MODELING OF RAILWAY SYSTEM MAINTENANCE AND AVAILABILITY BY MEANS OF COLORED PETRI NETS

MODELOWANIE UTRZYMANIA RUCHU I GOTOWOŚCI SYSTEMU KOLEJOWEGO ZA POMOCĄ KOLOROWYCH SIECI PETRIEGO

Prognostics and health management (PHM) technologies permit actionable information to enable proper decision-making for improving systems' performance. With the increasing requirements placed on the rail systems' availability, better maintenance decisions should be evaluated before practical application. The aim of this work is to build maintenance models and estimate the performance of considered maintenance decisions regarding the rail system's reliability and availability by means of Colored Petri nets. As a high-level formalization method, Colored Petri nets provide different color sets, which are suitable to represent different maintenance attributions. The maintenance models are evaluated at both the structure and parameterization levels. At the structure level, the structure correctness of the maintenance models is evaluated by using the state space analysis. At the parameterization level, specific maintenance decisions are illustrated. With various maintenance parameters, comparisons of system reliability and availability are made with the results obtained with the Colored Petri nets model.

Keywords: *prognostics and health management, colored Petri nets, railway system, maintenance, availability.*

Technologie prognostyki i zarządzania zdrowiem (PHM) dostarczają praktycznych danych, które umożliwiają podejmowanie właściwych decyzji w zakresie poprawy wydajności systemów. Wraz z rosnącymi wymaganiami dotyczącymi gotowości systemów kolejowych, rośnie potrzeba oceny decyzji dotyczących utrzymania ruchu przed ich wprowadzeniem w życie. Celem przedstawionej pracy było zbudowanie modeli utrzymania ruchu oraz oszacowanie za pomocą kolorowych sieci Petriego możliwości realizacji rozważanych decyzji konserwacyjnych dotyczących niezawodności i gotowości systemu kolejowego. Kolorowe sieci Petriego to metoda o wysokim poziomie formalizacji, którą w przedstawionej pracy wykorzystano do reprezentacji za pomocą różnych zestawów kolorów, różnych atrybutów utrzymania ruchu. Modele utrzymania ruchu oceniano zarówno na poziomie struktury jak i parametryzacji. Na poziomie struktury, poprawność struktury modeli utrzymania ruchu oceniano za pomocą analizy przestrzeni stanów. Na poziomie parametryzacji, zilustrowano konkretne decyzje dotyczące konserwacji. Niezawodność i gotowość systemu przy różnych parametrach utrzymania ruchu porównano z wynikami uzyskanymi za pomocą modelu kolorowych sieci Petriego.

Słowa kluczowe: *prognostyka i zarządzanie zdrowiem, kolorowe sieci Petriego, system kolejowy, utrzymanie ruchu, gotowość.*

1. Introduction

Maintenance plays an essential role in a system's life cycle. At the system level, the maintenance influences the reliability and availability of the system [3]. Achieving a high maintainability in the railway system requires a proper maintenance strategy. More maintenance means more life-cycle cost, while it may not lead to a dramatical improvement in the reliability. Hence, the performance of the strategy should be evaluated before it is put into the practical application.

Given that the enormous number of system components and maintainable items, it is a complex task to carry the analysis of railway system maintenance and availability. Railway system includes different subsystems, and the system structure will influence the overall system availability and performance. When managers plan a maintenance strategy, they have to take the system architecture into account. As in a free market, the optimal maintenance strategy can not only guarantee the availability of railway system but also have the best economic benefits. For the system maintenance and availability analysis, there are mathematical formulating and model-based analysis approaches. Garmabaki et al. presented the Multi-Attribute Utility Theory (MAUT), which used multiple objective functions to evaluate the cost and reliability of the maintenance optimization [5]. A gamma deterioration process was proposed by Meier-Hirmer et al., and it was applied to analyze the track maintenance [10]. Furthermore, the Maintenance Engineering Department of French National Railway Company (SNCF) introduced a formal method to estimate the maintenance strategy [1]. In publication [13], an application of stochastic Petri nets was presented to analyze the signal maintenance in France. All in all, comparing with the mathematical formulating approach, the model-based analysis can provide a more structured overview of the system. Additionally, it is much easier to read than the pure mathematical calculation [14].

Based on the existing database, some maintenance parameters are available. In order to evaluate the efficiency of the maintenance strategy, simulation-based analysis can be used to implement this task. Formalization & modeling can efficiently and cost-efficiently represent a real-world system. The system security analysis based on modeling is widely used in different research areas [17].

Stochastic-process techniques can be used to optimize the maintenance policies [4]. Due to the different overall system structure and

complicated mathematical calculation, in this case the simulation-based analysis is used as a modeling tool [14][8]. For instance, stochastic Petri nets are used to model the railway system maintenance and availability [4]. Reliability block diagrams and Monte Carlo simulation are applied to analyze different maintenance strategies on the product availability. As an alternative to reliability block diagram (RBD) and continuous-time Markov chain (CTMC) models, non-Markovian stochastic Petri nets have also been used to model maintenance processes of complex systems.

Formalization & modeling can efficiently and cost-efficiently represent a real-world system. The system security analysis based on modeling is widely used in different research areas [15]. The model-based safety analysis can be classified into two groups: (a) failure logic based. For instance, Fault Tree Analysis (FTA) emphasizes the model of failure propagation logic; (b) system states based. This approach addresses the analysis of the transition of system states, in order to identify the routes that a system transits from a safe state to a hazardous state [16].

However, these aforementioned methods are not really included in the availability analysis of an overall transportation system, which has to take different maintenance strategies into account. In particular, these methods cannot guarantee the correctness of analysis procedures, which will affect the conclusion. Hence, it is necessary to provide a methodology that can validate the model correctness when the maintenance decisions are evaluated, and permit to find out the best maintenance strategy.

Motivated by the problems mentioned above, the aim of this paper is to build maintenance models and estimate the performance of considered maintenance decisions regarding the rail system's reliability and availability by means of Colored Petri nets (CPN). This method takes full advantages of CPN to represent different maintenance attributions, validate the overall system's structure correctness, and evaluate the maintenance decisions.

In this paper, we build a CPN model which involves both the failure logic and system states to carry out the corrective and periodic maintenance strategies. This model can be reused for different subsystems. The subsystems are composing an overall system by using the system structure model, which can consider both parallel and series system structures. With the advantage of CPN, a color set that having different maintenance attributions is applied to represent the main parameters of the maintenance strategies [6]. What is more, CPN can also be used to verify the system functional safety, and more details can be found in our previous publication [14]. While this paper focuses on the analysis of system maintenance and availability by means of CPN, which validates the correctness of system structure and estimates the performance of considered maintenance decision in an overall system level.

The remaining of this paper is organized as follows: after the introduction, the modeling methodology is introduced in section 2. In this section, different maintenance strategies are taken into account during the modeling process. What is more, both the parallel and series system structures are involved, and the system architecture is verified based on state space analysis. Section 3 advances the methodology proposed in section 2 by illustrating a case study in a line of the Sweden railway. In this section, the CPN model and operation procedure are presented; the numerical simulation data with different maintenance strategies is available to evaluate the overall system availability. Finally, some conclusions are drawn in section 4.

2. Modeling methodology

The CPN model is used to represent the maintenance strategy and system structure. In a real system, each subsystem should provide a fault treatment process when it is out of service [11]. Different maintenance strategies, which can be verified according to the purpose to be achieved, are taken into account during the modeling process [3]. To simplify the expression of the CPN model, only nets and color sets structures are discussed, no arc expressions and binding elements are involved, and the guard functions are described by logical expressions.

2.1. Periodical maintenance of the system with different failure rates

Before the fault is treated, it may cause hazards. Maintenance methods can be categorized into the preventive treatment, corrective treatment, and condition-based treatment. In this paper, the condition-based treatment is not considered in the treatment components.

The following assumptions are considered when treatments are applied:

- as long as a system is failed, the corrective treatment will be activated;
- the downtime required for corrective maintenance time is set to follow an exponential distribution;
- the downtime required for the periodic maintenance is negligible;
- after the maintenance procedure, the component is as new as original;
- the failure behavior is stochastically independent.

Failure rates of components can be divided into three periods, as shown in Fig. 1. It includes early failures, random failures, and wear out failures. Each section is described with its Weibull distribution as $\lambda(t)=\frac{b}{T}\cdot\left(\frac{t}{T}\right)^{b-1}$, where b is the shape parameter of the failure slope, T is the characteristic lifetime, t is the service time. For electronic components, a burn-in process is required, it makes the components enter into the random failures period. At this period, the shape parameter is 1.0. Hence, the failure rates are constants for electronic components. However, $b>1$ indicates that the failure rate increases with time.

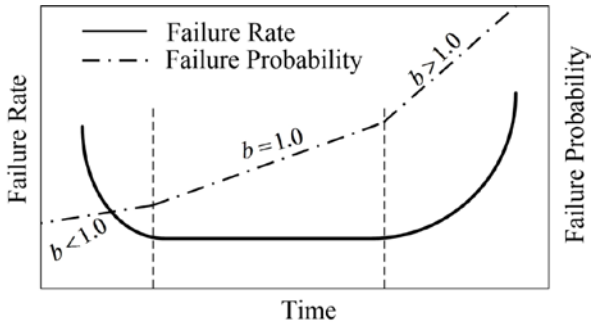


Fig. 1. Failure rates and Weibull parameter b

The probability density function of a Weibull random variable is given by:

$$f(t;T,b)=\frac{b}{T}\cdot\left(\frac{t}{T}\right)^{b-1}\cdot e^{-\left(\frac{t}{T}\right)^{b}} \tag{1}$$

where b is the shape parameter and T is the characteristic lifetime.

Failure probability:

$$F(t;T,b)=\int_{-\infty}^{t} f(t)dt=1-e^{-\left(\frac{t}{T}\right)^{b}} \tag{2}$$

Reliability:

$$R(t;T,b)=1-F(t)=e^{-\left(\frac{t}{T}\right)^{b}} \tag{3}$$

Failure rate:

$$\lambda(t)=\frac{Failures\_per\_Unit\_Time}{Quantity\_Exposed}=\frac{f(t;T,b)}{R(t;T,b)}=\frac{b}{T}\cdot\left(\frac{t}{T}\right)^{b-1} \tag{4}$$

For the numerical analysis, the system reliability is used to measure the probability that there is no failure happened before time t. The reliability of the system involved periodic treatment can be described as Eq. (5).

$$\begin{aligned}R_{PM}(t)&=\left[1-P\left(t\le k\cdot T_{PM}\right)\right]\cdot\left[1-P\left(k\cdot T_{PM}<t\le(k+1)\cdot T_{PM}\right)\right]\\&=\left[1-F\left(k\cdot T_{PM}\right)\right]\cdot\left[1-F\left(t-k\cdot T_{PM}\right)\right]=R\left(k\cdot T_{PM}\right)\cdot R\left(t-k\cdot T_{PM}\right)\end{aligned} \tag{5}$$

where T_{PM} is the periodic repairing interval; n means the n th periodic treatment; $R_{PM}(t)$ is the reliability of the system with the periodical treatment; $k\cdot T_{PM}\le t\le(k+1)T_{PM}$ is the time interval between the k th and $k+1$ treatment period; $k\in\mathrm{N}$. $P(T\le k\cdot T_{PM})$ indicates the probability that the system has failed in the period; $P\left(k\cdot T_{PM}<t\le(k+1)\cdot T_{PM}\right)$ is the failure probability after the k th renewal treatment. Hence, the reliability of the system with periodical maintenance is:

$$R_{PM}(t;T,b)=\exp\left[-\left(k\cdot\left(\frac{T_{PM}}{T}\right)^{b}+\left(\frac{t-k\cdot T_{PM}}{T}\right)^{b}\right)\right] \tag{6}$$

The reliability result is shown in Fig.2. When $b<1$, after the first time maintenance, the system reliability decreases when compared with the original reliability. When $b=1$, the periodical renewal does not influence the reliability performance of the system. When $b>1$, the R_{PR} is greater than the original reliability, it means that the system will have a higher functional operation probability with the periodical maintenance involving. Hence, only if the failure rate $\lambda'(t)>0$, the periodical maintenance comes with an advantage.

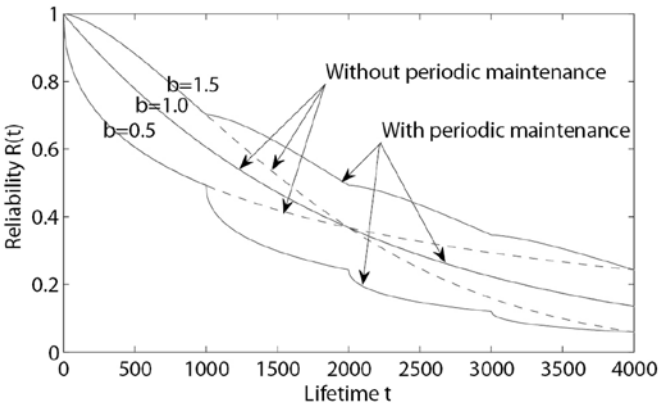


Fig. 2. System reliability with periodic maintenance

2.2. Maintenance and system architecture model

A maintained subsystem consists of these attributes: name, state, service time with relevant parameters, operation time, maintenance period, and characteristic lifetime. The subsystem has a name to identify it, and it can be in either fail or operating state. When the subsystem is failed, it will be repaired in a specific time. Section 2.1 illustrates a system's reliability with involving a periodic maintenance. Given that Weibull distribution is widely applied in the practical engineering, it can describe many behaviors of a failed system [1]. When we apply the Weibull distribution to predict the subsystem's service time, the characteristic lifetime T, service time τ, and shape parameter b should be involved. Additionally, by changing the shape parameter b, the Weibull distribution can be easily adapted to different system characteristics.

To represent the subsystem's aforementioned attributions, the declarations of the color structures are presented as in Fig. 3. Here, color sets *component* and *Data* represent subsystems and operation data, respectively. The subsystem with timestamps have the following attributions in *component*: subsystem name (*subsystem*), timestamps (*timestamp*), service time τ (*servicetime*), subsystem state (*state*), periodical maintenance period T_{PM} (*maintenance period*), characteristic lifetime T (*characteristic lifetime*), shape parameter *b* (*shape*), and mean time to repair (*MTTR*).

```
colset subsystem=string;
colset timestamp=int;
colset servicetime=int;
colset state=with operate|fail;
colset maintenance period=real;
colset characteristic lifetime=real;
colset shape parameter=real;
colset MTTR=real;
colset component=with
subsystem|timestamp|servicetime|state|maintenance
period|characteristic lifetime|shape parameter|MTTR timed;
colset Data=with subsystem|timestamp|state;
```

Fig. 3. Color set declaration

Fig. 4 shows the net's structure of the CPN model. The sub-nets of the system structure share the same model layout but with different logics. For a parallel system, the overall system is failed only if all the components are failed, as shown in Eq. (7). As long as one component is repaired, the overall system will operate again, as shown in Eq. (8). Similarly, the logical expressions for a series system's failure and recovery are described by Eq. (9) and (10), respectively.

$$\neg(\#state:A)\wedge\neg(\#state:B) \tag{7}$$

$$(\#state:A)\vee(\#state:B) \tag{8}$$

$$\neg(\#state:A)\vee\neg(\#state:B) \tag{9}$$

$$(\#state:A)\wedge(\#state:B) \tag{10}$$

where $\neg$ means "not"; $\#state:A$ indicates the state of component A; $\wedge$ is "and"; $\vee$ is "or"; the transition will be enabled when the logical expression result is *true*.

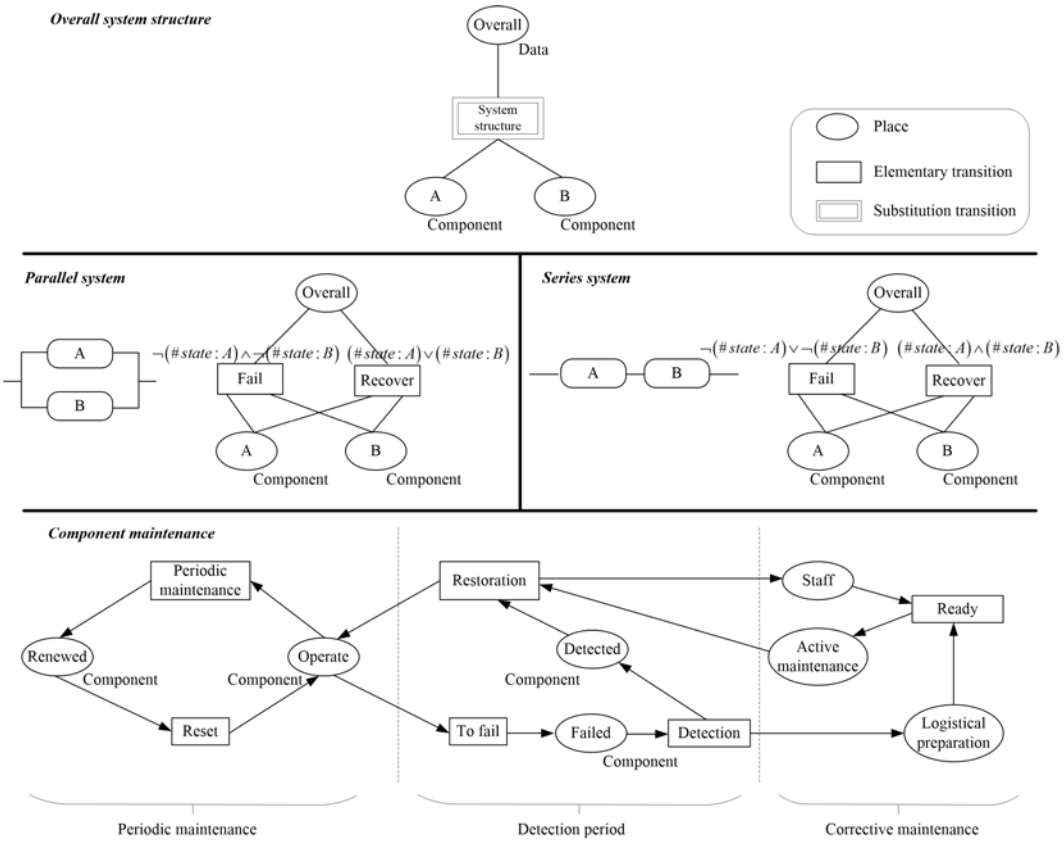


Fig. 4. CPN module of system

The component maintenance involves periodic maintenance, detection period, and corrective maintenance. The transition *Periodic maintenance* will be executed automatically based on the maintenance period T_{PM}. The transition *To fail* will be fired if component A is out of service before the next time periodical maintenance. After the failure is detected and the preparation and staff are ready, the transition *Restoration* will be enabled to carry the maintenance process, and reset the component state from *fail* to *operate*.

The architecture verification is carried out by using the model checking method [14]. To implement this verification, the state space analysis is required. The formal description of a system failure happening can be treated as that: in a marking M_i, when an enabled binding element (t,b) occurs, it will change the marking M_i to another marking M_{i+1}, defined by:

$$\forall p \in P: M_{i+1}(p) = M_i(p) - \sum_{(t,b)\in Y} E(p,t)<b> + \sum_{(t,b)\in Y} E(t,p)<b> \quad (11)$$

where $M_i(p)$ is the number of $M_i(p) - \sum_{(t,b)\in Y} E(p,t)<b>$ removes tokens from M_i while $\sum_{(t,b)\in Y} E(t,p)<b>$ added tokens to M_{i+1}. Moreover, M_{i+1} is directly reachable from M_i by the occurrence of the step Y, which denote as $M_i[Y > M_{i+1}$.

Take a parallel system that has two components *A* and *B* as an example, both *A* and *B* are failed, the overall system is failed. Here, M_i represents that both *A* and *B* are failed; step Y represents the transition *Fail*; M_{i+1} means the *Overall* is failed. Markings M_i and M_{i+1} are found through the function *Check*(). In order to limit the size of state space, other parameters of the subsystem are set to 0 except for the name and state. This attribution can be checked by the following query.

```
---------Check query---------
fun  Check()=PredAllNode(fn  n=>Mark.Parallel.Overall
1 n=[("NULL",operate)]
andalso Mark.Parallel.A 1 n=[("A",0,0,fail,0,0.0,0.0)]
andalso Mark.Parallel.B 1 n=[("B",0,0,fail,0,0.0,0.0)]);
val M_i=Check();
OutArcs(p)
---------Results---------
val M_i=[p]:Node list
val it=[q]:Arc list
---------Check query---------
val Y =ArcToTI(p);
val M_i+1=DestNode(p);
Mark.Parallel'Overall 1 Mi+1;
Mark.Parallel'A 1 Mi+1;
Mark.Parallel'B 1 Mi+1;
---------Results---------
val Y =Parallel'Fail 1:TI.TransInst
val M_i+1 =r:Node
val it=[(„AB",fail)]
val it=[(„A",0,0,fail,0,0.0,0.0)]
val it=[(„B",0,0,fail,0,0.0,0.0)]
```

As shown in the results, M_i is found in the node p, which represents that both components *A* and *B* are failed. In addition, arc q is the only out arc of M_i, and its destination marking is M_{i+1} in the node r. The binding transition of arc q is *Parallel'Fail*. The results mean the parallel system architecture meets the requirement.

3. Case study

3.1. System description

Railway involves different blocks, which are the minimum operation sections in the railway transportation. Fig. 5 is the RBD for a block section. A block is composed of the following subsystems [12].

- Track circuit: is used to detect the train location on the trail track.
- Interlocking: prevents conflicting movements by receiving information from other subsystems, outputs movement restrictions to make sure that the train can operate safely. It can be categorized as mechanical, electro-mechanical, relay, and electronic interlocking.
- Radio Block Center (RBC): is in charge of analyzing the train's position and transferring it to the interlocking system, and sending the movement authority and other commands to the train via radio.
- Signal: gives the driver pass information based on the state of the line ahead. It can be a mechanical or electrical device.

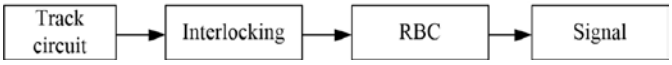


Fig.5. RBD of a block section in full functional operation

Table 1. Parameters for MTTF, MTTR, and maintenance strategy

Subsystem	MTTF (year)	MTTR (h)	Initial strategy	Shape b
Signal	2.1464	5.14	Corrective maintenance	1.0
Interlocking	2.8581	5.54	Corrective maintenance	1.5
Radio block center	2.8581	5.14	Corrective maintenance	1.0
Track circuit	2.004	3.36	Corrective maintenance	1.0

The maintenance data for this model is selected from a 203 km long railway in Sweden [12]. Note that the data is only used to do the illustration and show the efficiency of our approach. The results can be various based on different actual data. As shown in Table 1, the mean time to failure (MTTF), mean time to repair (MTTR), maintenance strategy, and distribution shape b are given. Certain assumptions are made that the signal, RBC, and track circuit subsystems are consisted of electrical items, and they are supposed to have constant failure rates. Hence, the distribution shape is $b = 1$. However, the interlocking subsystem involves not only electrical but also mechanical elements, it is assumed that the mechanical wear should be taken into consideration. Hence, the distribution shape is $b = 1.5$.

3.2. CPN model

Based on the aforementioned modeling methodology, the RBD in Fig. 5 can be transferred into the CPN model as shown in Fig. 6.

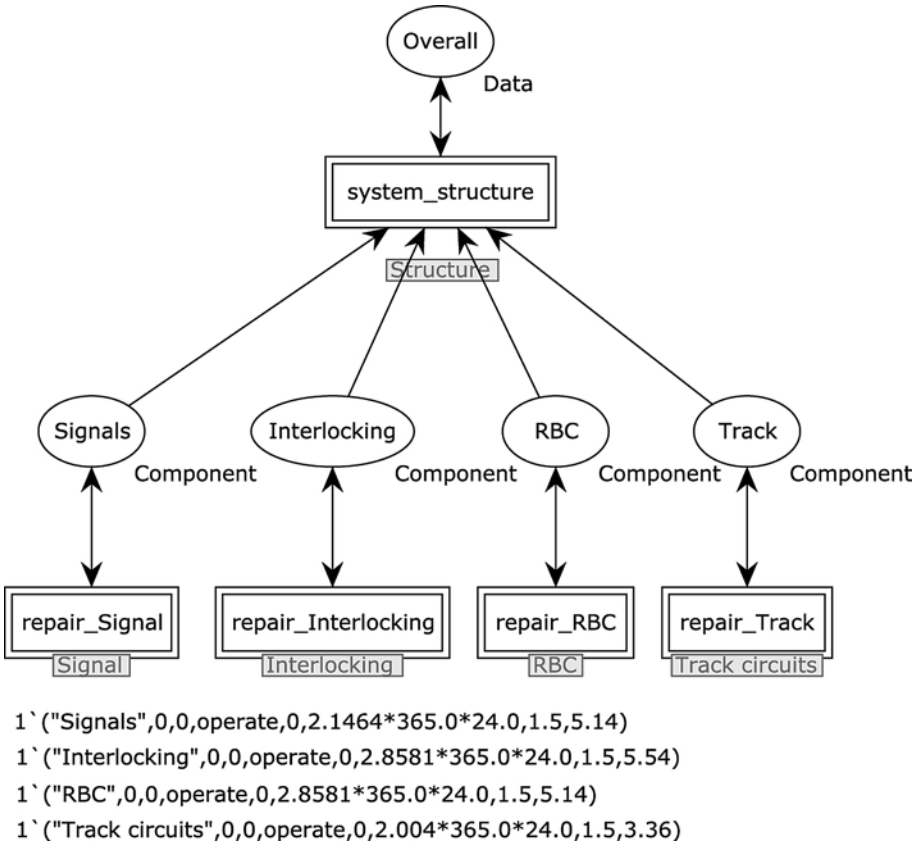


Fig. 6. CPN model of system structure

Each component is assigned with parameters based on the maintenance strategy in Table 1. The subnet *system_structure* represents the system architecture, which determines the rules to evaluate the overall system performance. In order to evaluate the full functional operation performance, which means all these components should operate correctly, these four subsystems are involved as the series system structure in this model.

The operation procedure of the CPN model can be shown in Fig. 7. Each subsystem, which has the attributions as shown in the color set *component*, has the periodic and corrective maintenance strategies. If the subsystem is failed before the periodic maintenance, the corrective maintenance is activated and the subsystem will be as new as original. The *Component* will be updated in real time, and once the binding element of transition *Fail* or *Recover* is enabled, the token in the color set *Data* will be updated. By monitoring the token on place *Overall*, the system performance can be recorded. Color set *Data* can record: which subsystem failed and caused the overall system out of service; the timestamps of when the system failed or recovered. For instance, when the subsystem signal is out of service at 3858 hours, there will one token (*1`("Signals",3858,fail)*) on the place *Overall*; when the subsystem is repaired in 2 hours later, the token will be updated to *1`("NULL",3860,operate)*.

3.3. Simulation data analysis

Note that, the train operation will be not totally out of service when one component failed. For instance, when the signal is out of service, the railway operation can still be possible if the driver obtains the permission from dispatchers and drives the train with a speed less than 40 km/h in the visual supervision mode. Hence, different sorts

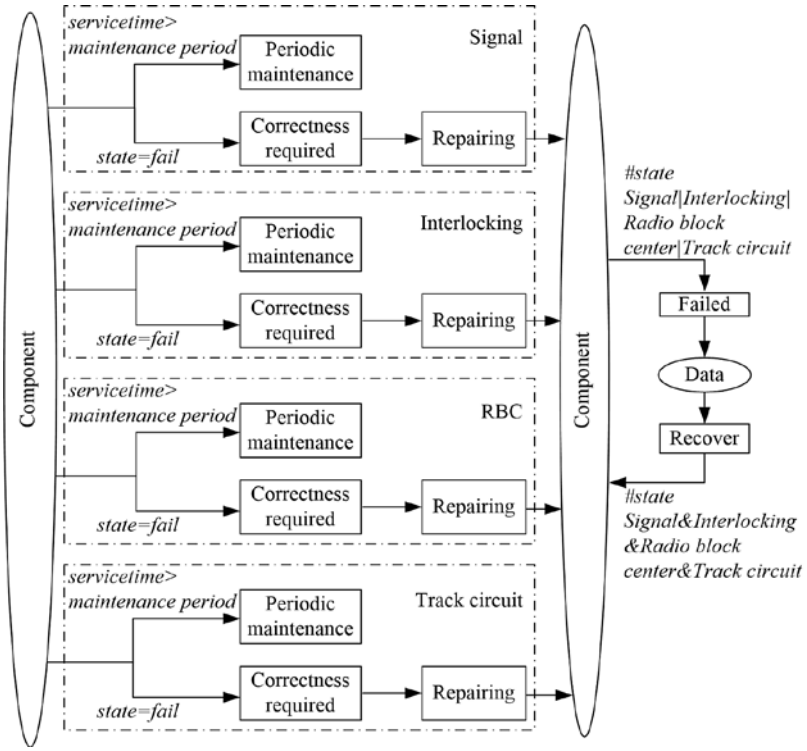


Fig. 7. CPN operation procedure

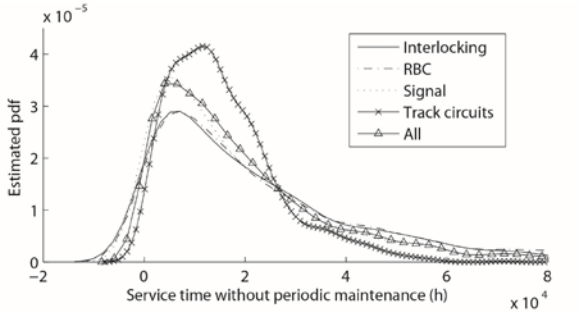


Fig. 8. Probability density function of the service time without periodic maintenance

of overall availabilities can be carried by determining respective rules.

Reliability ($R(t)$) is defined as the probability that a component performs during an operating time interval t. The achieved availability (A_A) can be obtained by calculating the mean time between maintenance ($MTBM$) actions and the average downtime $\bar{M}$ [7].

$$R(t) = P(T > t) \tag{12}$$

$$A_A = \frac{MTBM}{MTBM + \bar{M}} \tag{13}$$

Without having an onerous computation, the system reliability and availability calculation can be carried out by the statistical analysis of the CPN model simulation results.

As monitoring the *Overall* system, the performances of the subsystems can be monitored as well. Before the periodic maintenance is involved, the service time of each subsystem and the overall system is shown in Fig.8. It indicates that the track circuits' subsystem has the most considerable influence on the overall system service time.

By involving a periodic maintenance procedure to the initial maintenance strategy for each subsystem, the difference in the service time can be available to do the analysis. The periodic maintenance interval is 2000 hours. As shown in Fig.9, the distributions of the track circuits and overall system are changed dramatiscally when involving the periodic maintenance. However, the distributions of the interlocking, RBC, and signal subsystems are not changed. This result is also corresponding to the mathematical calculation conclusion in the section 2.1: with a different shape parameter b the periodical renewal has the respective influence on the system performance.

As shown in Fig.10, the reliability of the overall system is increased with the periodic maintenance involving. Fig. 11 indicates the relation between the availability and the periodic maintenance interval. Given that more periodic maintenance means more cost, the

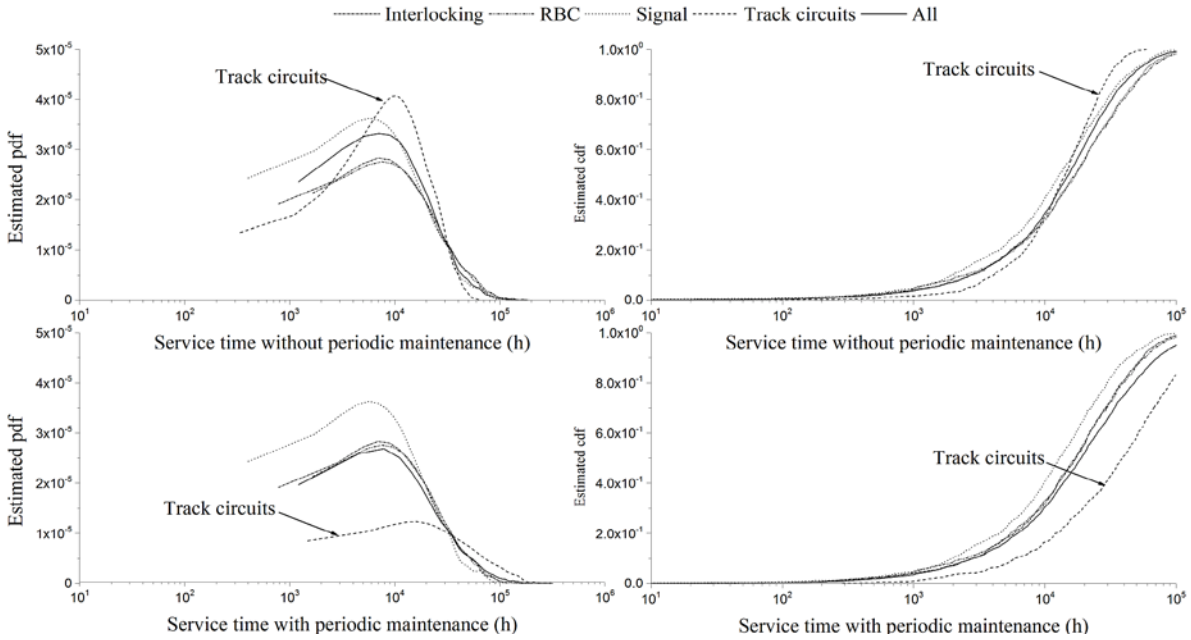


Fig. 9. The service time with or without a periodic maintenance

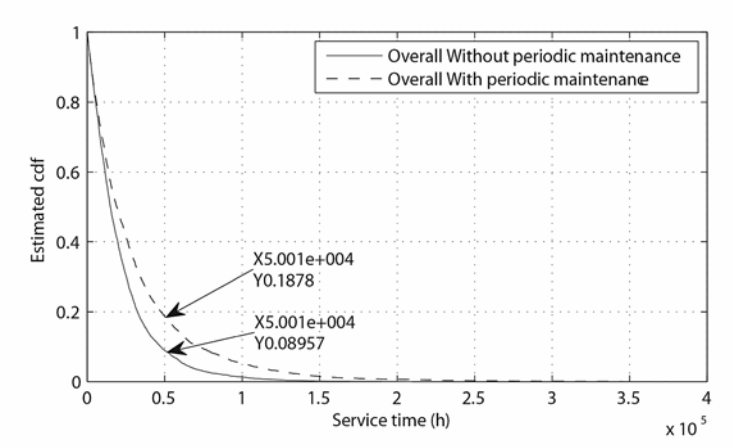


Fig.10. Overall system reliability

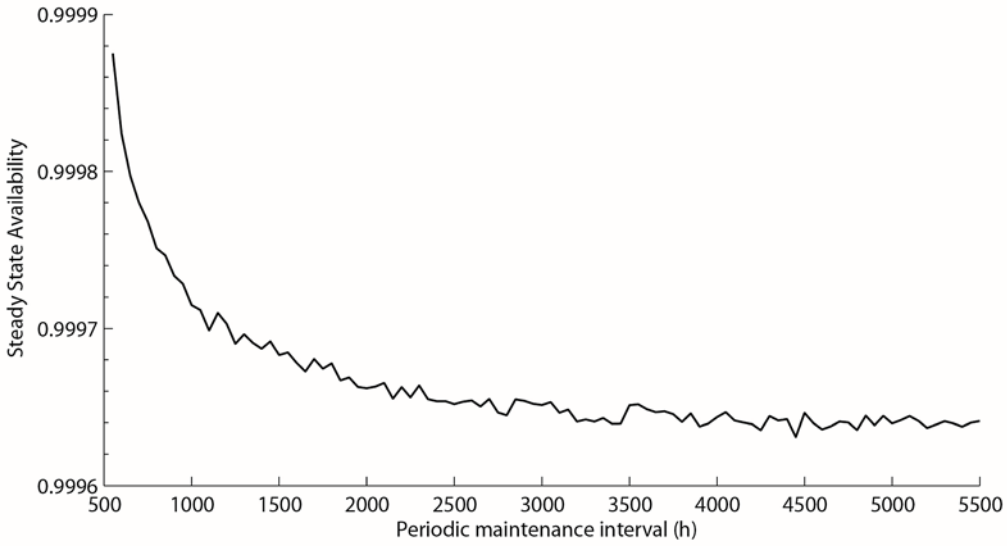


Fig. 11. Relation between the availability and the periodic maintenance interval

maintenance strategies can be modified based on the reliability and availability target.

4. Conclusion

In this paper, a maintenance CPN model, which can implement both corrective and periodical maintenances, was proposed to represent the maintenance procedure. Different subsystems were connected to compose an overall system by the system structure model, which can deal with both parallel and series systems. With the high expressive color sets in the CPN, different maintenance attributions were involved. The maintenance parameters can be modified easily to meet the practical requirements.

The results indicated that using the CPN model to simulate and analyze the overall system reliability and availability will be more efficient than the mathematical calculation. The structure of the overall system was represented by the subnet *system structure*, which was controlled by the logical expression in the CPN model. All in all, with the assistant of CPN model, the maintenance strategy can be evaluated before putting into practice, and the model permitted to find out the best maintenance strategy.

Acknowledgement:
I hereby present my gratitude to the China Scholarship Council and the iVA, Technische Universität Braunschweig, Germany. This paper is supported by the State Key Laboratory of Rail Traffic Control and Safety (Contract No.RCS2017K005), Natural Science Foundation of China (U1734211). We would also like to thank the Editors and all anonymous reviewers for their valuable comments. Thanks to our colleague Susanne Arndt for her language modification.

References

1. Antoni M, Ammad N. Formal validation method and tools for Frenchcomputerized railway interlocking systems, 2008: 99-106.
2. Bertsche B. Reliability in automotive and mechanical engineering: determination of component and system reliability. Springer Science & Business Media, 2008.
3. Birolini A. Reliability engineering. Heidelberg: Springer, 2007.
4. Dersin P, Valenzuela R C. Application of non-Markovian stochastic Petri nets to the modeling of rail system maintenance and availability, Simulation Conference (WSC), Proceedings of the 2012 Winter. IEEE, 2012: 1-12.
5. Garmabaki A H S, Ahmadi A, Ahmadi M. Maintenance optimization using multi-attribute utility theory. Current Trends in Reliability, Availability, Maintainability and Safety. Springer, Cham 2016: 13-25, https://doi.org/10.1007/978-3-319-23597-4_2.
6. Jensen K, Kristensen L M. Coloured Petri nets: modelling and validation of concurrent systems. Springer Science & Business Media, 2009, https://doi.org/10.1007/b95112.
7. Katukoori V K. Standardizing availability definition. University of New Orleans, New Orleans, La., USA, 1995.
8. Kowalski M, Magott J, Nowakowski T, et al. Analysis of transportation system with the use of Petri nets. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2011: 48-62.
9. 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.

10. Meier-Hirmer C, Riboulet G, Sourget F, et al. Maintenance optimization for a system with a gamma deterioration process and intervention delay: application to track maintenance. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2009, 223(3): 189-198, https://doi.org/10.1243/1748006XJRR234.
11. Miyagi P E, Riascos L A M. Modeling and analysis of fault-tolerant systems for machining operations based on Petri nets. Control Engineering Practice 2006, 14(4): 397-408, https://doi.org/10.1016/j.conengprac.2005.02.002.
12. Morant A, Gustafson A, Söderholm P. Safety and availability evaluation of railway signalling systems, Current Trends in Reliability, Availability, Maintainability and Safety. Springer International Publishing 2016: 303-316, https://doi.org/10.1007/978-3-319-23597-4_22.
13. Quiroga L, Slovak R, Wegele S, et al. Petri nets based simulation and optimisation of railway maintenance processes. FORMS/FORMAT, 2008.
14. Song H, Liu J, Schnieder E. Validation, verification and evaluation of a Train to Train Distance Measurement System by means of Colored Petri Nets. Reliability Engineering & System Safety 2017, 164: 10-23, https://doi.org/10.1016/j.ress.2017.03.001.
15. Schnieder E, Schnieder L, Müller J R. Conceptual foundation of dependable systems modelling[J]. IFAC Proceedings Volumes 2009, 42(5): 198-202, https://doi.org/10.3182/20090610-3-IT-4004.00039.
16. Zhao L, Thulasiraman K, Ge X, et al. Failure Propagation Modeling and Analysis via System Interfaces. Mathematical Problems in Engineering 2016, 2016.
17. Zhang T, Cheng Z, Liu Y, et al. Maintenance scheduling for multi-unit system: a stochastic Petri-net and genetic algorithm based approach. Eksploatacja I Niezawodnosc- Maintenance and Reliability 2012 (3): 256-264.

Haifeng SONG
Eckehard SCHNIEDER
Institute for Traffic Safety and Automation Engineering, Braunschweig
Technische Universität Braunschweig
Hermann-Blenk-Straße 42
38108 Braunschweig, Germany

E-mail: h.song@tu-bs.de , e.schnieder@tu-bs.de

Article citation info:
YU H, ZHANG G, RAN Y, LI M, WANG Y. A comprehensive and practical reliability allocation method considering failure effects and reliability costs. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 244–251, http://dx.doi.org/10.17531/ein.2018.2.09.

Hui YU
Genbao ZHANG
Yan RAN
Mengqi LI
Yang WANG

A COMPREHENSIVE AND PRACTICAL RELIABILITY ALLOCATION METHOD CONSIDERING FAILURE EFFECTS AND RELIABILITY COSTS

KOMPLEKSOWA I PRAKTYCZNA METODA ALOKACJI NIEZAWODNOŚCI UWZGLĘDNIAJĄCA SKUTKI USZKODZEŃ I KOSZTY NIEZAWODNOŚCI

In view of the drawbacks in existing allocation methods which are incomplete considerations and poor practicality, a comprehensive fuzzy allocation method considering failure effects and reliability costs is proposed. Fuzzy linguistics and triangular fuzzy numbers are used to evaluate the uncertainty and subjective factors in allocation process. The traditional risk priority numbers (RPNs) are modified to overcome the shortages which are the same factor weights and equal difference of failure effects in original methods. State of the arts, components intricacy and working conditions are used to construct the reliability costs model, which solves the difficulties of costs statistics and avoids the sophisticated calculations which exist in current allocation methods. The relationship between reliability costs and potential risk of subsystem is studied and the value range of it is given in this paper. A case example is given to illustrative the scientificity and practicability of proposed allocation method.

Keywords: *reliability allocation, reliability cost, failure modes and effect analysis (FMEA), relative reliability, fuzzy methods.*

Ze względu na niedostatki istniejących metod alokacji, które nie dają pełnego obrazu problematyki i mają słabe zastosowanie w praktyce, w artykule zaproponowano kompleksową metodę alokacji opartą na logice rozmytej, uwzględniającą skutki uszkodzeń i koszty niezawodności. W pracy wykorzystano lingwistykę rozmytą i trójkątne liczby rozmyte do oceny niepewności i czynników subiektywnych w procesie alokacji. Zmodyfikowano tradycyjny wskaźnik liczby priorytetowej ryzyka (RPN), co pozwoliło na poprawę mankamentów charakteryzujących oryginalną metodę, t.j. takie same współczynniki wagowe i równoważność skutków uszkodzeń o różnym stopniu ciężkości. Na podstawie wiedzy o stanie techniki, złożoności komponentów i warunkach pracy, skonstruowano model kosztów niezawodności, który rozwiązuje trudności dotyczące sporządzania statystyki kosztów i pozwala uniknąć skomplikowanych obliczeń stosowanych w obecnych metodach alokacji. Zbadano związek między kosztami niezawodności a potencjalnym ryzykiem podsystemu, oraz podano jego zakres wartości. Prezentowane studium przypadku demonstruje możliwe zastosowania i efektywność proponowanej metody.

Słowa kluczowe: *alokacja niezawodności; koszty niezawodności; analiza przyczyn i skutków uszkodzenia (FMEA); względna niezawodność; metody rozmyte.*

1. Introduction

Reliability allocation is a vital step of reliability design. A scientific allocation method can make the system owns the highest reliability while expending the minimum costs. The current allocation methods including traditional methods, risk priority number based (RPN-based) methods and cost-based methods etc. The above-mentioned allocation methods have following shortages ubiquitously:

(1) The traditional allocation methods do not consider the failure effect on system, which makes the results incredible.
(2) The same weight of factors and equivalence relationship between different severities cause the results of RPN-based methods deviating from reality.
(3) It is difficult to obtain the concrete cost statistics, and the calculation process of cost function is too complex, which makes the cost-based allocation methods impractical.

In addition, the current allocation methods just allocate from single aspect, either considering the failure effects or manufacturing costs, which are lack of a comprehensive consideration. Thus, it is difficult to optimize the allocation results.

Therefore, this paper proposes a comprehensive allocation method considering both failure effects and reliability costs. The risk priority numbers in current RPN-based methods are modified to represent the failure effects. Based on the inspiration of generalized cost function, the reliability costs are described by the current and the highest reliability of system. State of the art, working conditions and subsystem intricacy are considered synthetically to construct a semi-quantitative cost function. Value range of the relationship between the costs and potential risk of subsystem is given to avoid the smaller weight factor is too small to take into consideration, which ensures the scientificity of the results.

2. Literature review

Reliability allocation is to distribute the reliability target of system to its component subsystems actually through a specific method. It must satisfy the reliability requirements of system but also a variety of constrains. The most basically requirement is to solve the following inequality:

$$\begin{cases} R_s(R_1, R_2, \cdots, R_i, \cdots, R_n) \geq R_s^* \\ \vec{g}_s(R_1, R_2, \cdots, R_i, \cdots, R_n) \leq \vec{g}_s^* \end{cases}. \quad (1)$$

Where R_s^* is the reliability target of system, $\vec{g}_s^*$ is the constrain of system which includes cost, volume and mass factors etc.., R_i is the reliability target of subsystem i.

Assuming that a serial system is composed by k subsystems, $\lambda^*(t)$ is the target failure rate of system. $\lambda^*_i(t)$ is the allocated failure rate of subsystem i which can be expressed as:

$$\lambda^*_i(t) = \omega_i \cdot \lambda^*(t) \ , \quad t \geq 0 \ , \ i=1,2,\ldots,k. \quad (2)$$

where ω_i is the allocation weight of subsystem i, which can be obtained by the following equation:

$$\omega_i = \frac{n_i}{\sum_{i=1}^{k} n_i} \ , \quad i=1,2,\ldots,k. \quad (3)$$

where n_i is the estimated value of subsystem i, it could be component numbers in subsystems or failure rate or others [9]. The various allocation methods in the end are the different selection of n_i.

2.1. Traditional allocation methods

Traditional reliability allocation is a method that considers single or multiple factors, judge subsystems by objective or subjective information and finally calculates allocation weight of subsystems through a certain of combination operations. It is aimed at guiding new designs by the current reliability level of systems, that is, the higher reliability the existing subsystem, the lower failure rate the corresponding new subsystem allocated.

State of the art, intricacy, operating time and working conditions of system are closely related to its reliability level, therefore, these factors are always regarded as the consideration factor while lacking of reliability data. Many scholars utilized various operations to allocate subsystems weight by the consideration of above four factors.
or summation [14]:

$$n_i = \sum_{j=1}^{4} A_{ij} \ , \ i=1,2,\ldots,k. \quad (4)$$

where A_{ij} is the estimated value of factor j for subsystem i, which value rang is the natural number from 1 to 10.
or multiplication [7]:

$$n_i = \prod_{j=1}^{4} A_{ij} \ , \ i=1,2,\ldots,k. \quad (5)$$

or mixed operations [1]:

$$n_i = A_{i1}(A_{i2} + A_{i3} + A_{i4}) \ , \ i=1,2,\ldots,k. \quad (6)$$

where A_{i1} is the state of the art of subsystem i.

Karmiol [13] allocated the reliability index by Eq.(4) while considering the state of the art, intricacy, criticality and operating time as evaluating factors.

To solve the problem of the same weight between judging factors, the evaluation results are modified by factor weights or expert weights in some papers[19,27]. O'Hagan[20] presents a calculating method of relative weight a_j by maximal entropy, the estimated value n_i of subsystem i is given as:

$$n_i = \sum_{j=1}^{n} a_j A_{ij} \ , \ i=1,2,\ldots,k. \quad (7)$$

Wang et al.[23] take the failure frequency, failure severity, subsystems maintainability and complexity etc. seven factors into account, evaluate the allocated value of subsystems by Eq. (7) after seven factors were compared each other by both quantitative and qualitative information. Where a_j is the relative weight of factor j to others. A_{ij} is the relative value of subsystem i to subsystem j.

Though the traditional allocation methods can works in a certain extent in system allocation, these methods do not take the failure effects into consideration, nor take the manufacturing costs of system into consideration.

2.2. RPN-based allocation methods

It is inevitable for any systems to have no failure during it runtime. Various failures bring different influences to system, even if the same failure mode occurs in different subsystems. Whatever the failure happens, it would cause a loss to system more or less. Therefore, it must take the potential failures and failure effects into consideration while the reliability of system is allocated.

Recently, some scholars [11, 26] proposed the RPN-based allocation methods. RPN is the scale of failure criticality, measuring the severity(S), occurrence (O) and detection (D) though an ordinal scales from 1 to 10 in the failure modes and effects analysis (FMEA) of system. The RPN of failure mode j in subsystem i as given below:

$$RPN_{ij} = S_{ij} \times O_{ij} \times D_{ij} \ , \quad (8)$$

When the detection is considered in the severity of failures[5,11], the Eq. (8) can be rewritten as below:

$$RPN_{ij} = S_{ij} \times O_{ij} \ , \quad (9)$$

Assuming that there is N failure modes in the system, Itabashi-Campbell [11] proposed the estimated value of subsystem i can be given by Eq.(10) or Eq.(11) according to the different intentions of allocators.

$$n_i = B_i \ , \quad (10)$$

$$n_i = 1 - \frac{B_i}{\sum_{i=1}^{k} B_i} \ , \quad (11)$$

where:

$$B_i = \frac{1}{N}\sum_{j=1}^{N} S_{ij} \times O_{ij} \ , \tag{12}$$

Many researchers [18,24] point out it is unreasonable to give the same weight to risk factors, for instance, the failure mode S_1=2, O_1=8 and S_2=8, O_2=2 has the same RPN in this manner though it was not the case in reality.

To overcome the defects of this method, a new allocation approach was presented by Kim et al. [15]. The original severity is modified by an exponential function, assuming that S_{ij} is the original severity of failure mode j in subsystem i, the new severity is given as :

$$\tilde{S}_{ij} = \exp(\alpha S_{ij}) \ , \tag{13}$$

where a is the severity coefficient, which is depended on the designer intention. a must get a higher value while the designer take failure effects more seriously and vice versa.

The evaluation criterion of subsystem i is given by:

$$n_i = \frac{1}{m_i \tilde{S}_i F_i} \ , \tag{14}$$

where:

$$\tilde{S}_i = \max(\tilde{S}_{i1}, \tilde{S}_{i2}, ..., \tilde{S}_{iN_i}) \ , \tag{15}$$

$$j_i = \arg\max_j \tilde{S}_{ij} \ , \tag{16}$$

m_i is the number of failure mode which having the same severity with $\tilde{S}_i$. F_i is the frequency ratio of failure mode j_i in subsystem i.

Though this method solves the shortcoming of equal weighed in general RPN-based methods, it is still unreasonable. The evaluation values in the paper are specific numbers which is far away from the actual due to the subjectivity and uncertainty in judgment processes [2,24]. Furthermore, whatever the RPN-based allocation methods only consider the failure effects on system, and ignore the necessary manufacturing costs of system with a specific reliability during it produced.

2.3. Cost-based allocation methods

Generally speaking, everybody wants the system with higher reliability, but the higher reliability of system, the more manufacturing costs needed, and sometimes even lose more than gained. Therefore, the manufacturing cost is the essential factor that must be taken into accounts in any systems development.

The current cost-based allocation methods mainly focused on the optimal planning of allocation, there are two main ways to consider the costs. The one is regarding costs as a specific constant which obtained from statistics or assumption, the other is considering the cost as an increasing function with the reliability of system[3,8,10,12, 21].

Todinov [22] regards the costs and losses of system as the consideration factors in allocation. Assuming Q_i is the manufacture costs of subsystem i, the losses caused by failures of subsystems is the constant L, the total costs C_i of subsystems as given below:

$$C_i = Q_i + L \ , \tag{17}$$

Wang et al. [23] weight the costs through cost sensitivity, which is obtained from the experts experience by a scale of 0 to 1 value to represents the relationship between the costs and reliability of subsystem i. :

$$C_i = \frac{\Delta C_i}{\Delta R_i} \ , \tag{18}$$

where ΔC_i is the increased costs of subsystem i. ΔR_i is the improved reliability of subsystem i.

In the actual project, however, the cost of systems is hard to collect with the changing of technological and price level. In addition, it is unreasonable to treat the costs of various failures as a constant while there is a big difference effects between all kinds of failure modes.

Dale et al. [4] proposed the six basic properties of cost function in 1986, regarding costs as the increasing function with reliabilities. And then many scholars set up cost function model on this basis. Based on the six properties, Li et al. [17] establish the cost function of diesel engine as:

$$c(R_i) = f_i \ln \frac{R_{i,\max} - R_{i,\min}}{R_{i,\max} - R_i} \ , \tag{19}$$

where R_i is the allocated reliability of subsystem i. f_i is the cost coefficient of subsystem i, where $0<f_i<1$. $R_{i,\max}$ and $R_{i,\min}$ is the maximum reliability under the current technologies and the current reliability of subsystem i respectively. $c(R_i)$ is the improvement costs of subsystem i from the reliability $R_{i,\min}$ to R_i.

According to the three properties mentioned in [16] that a cost function must be a positive definite function and non-decreasing and increasing rapidly as reliability close to 1, Elegbede [6] presents the total costs, which is expressed by:

$$C_s = \sum_{i=1}^{s}\sum_{j=1}^{k_i} k_i \bullet h_i\left(\frac{\log(1-R_i)}{k_i}\right) \ , \tag{20}$$

where k_i is the number of components in subsystem i. R_i is the reliability of subsystem i. s is the number of subsystems and $h_i(\cdot)$ is the function with the three properties.

Though cost function could describe the relationship between the costs and reliability of subsystems in a certain extent, it is poor practicability in the practical application due to its complicated computing processes.

Recently, Yadav et al. [25] notice the efforts of reliability improvement and describe it as a function which related to failure rate. The modified evaluation criteria of subsystem i based on the method presented by Kim is given by:

$$n_i = \frac{m_i \tilde{s}_i}{\delta_i e_i} \ , \tag{21}$$

where δ_i is the difficulty coefficient of subsystem i for improvement. e_i is the effort coefficient where $e_i = \ln\lambda_i / \sum_{i=1}^{k} \ln\lambda_i$.

Though the approach presented by Yadav noticed the influence between the reliability of subsystems and improvement efforts, it ig-

nored the real determinant factor of efforts is the current technology level of subsystems rather than failure rate. After the severities and efforts are modified, each subsystem is multiplied with different degree of difficulty coefficients according to the subjective consciousness of the allocators, which is equivalent to modify the efforts twice. It is no doubt to increase the subjectivity in the allocation process, resulting in lower credibility of the distribution results.

The deficiencies stated above urgently require a more thoughtful and credible allocation method.

3. Proposed allocation method

Aimed at these defects mentioned above, we present the solution in this section. Fuzzy linguistic is used to describe the uncertainty subjective information in the allocation process. A more practical reliability allocation method which integrates failure modes and the necessary manufacture cost of system in a specific reliability is proposed. The steps and basis of proposed approach are shown in detail as following.

Step 1 Influential factors determination

The effect on system caused by the failures of components is defined as the potential risk (*PR*) of subsystems. Any system is made up of several subsystems, and there are several potential failure modes in each subsystem. The potential risk is determined by both severity and occurrence of failure mode in subsystems. Therefore, the *S* and *O* must be considered in the allocation process. Secondly, the improvement of reliability in any systems must increase the manufacture costs. Everyone expects that the system owns a higher reliability, but it is always not the case due to the constraint of costs even if the existing technological level could achieve. Due to these reasons, the corresponding costs (*C*) of system in a specific reliability must be considered.

The precise costs data of system is hard to collect, and it is not feasibility for various products even the data has gotten. Allocation methods based on cost function are too complicated in computing process to be applied in practical application. From the previous researches (as mentioned before), the necessary costs of reliability improvement is constrained by both the current reliability level and the highest reliability level under the circumstance of subsystem. And the reliability of systems is closely associated with the state of the art (*SA*), subsystem intricacy (*SI*), operating time (*OT*) and environmental conditions (*EC*). For these reasons, we set the four factors as the related factors of the corresponding manufacturing costs of system in a specific reliability. To simplify the allocation process, the operating time is ignored for the reason that it is same in a system even under different technological levels. Finally, the influence set *K* is expressed as

$K=\{PR, C\}=\{(O,S), (SA, SI, EC)\}=\{$(occurrence, severity), (state of the art, intricacy, environmental condition)$\}$

Step 2 Experts rating

Experts are asked to rate the influence factor set K on the basis of objective information and subjective judgment. Since mainly of the collected data are incomplete or imprecise, also the opinions of design makers are essentially vague, information description using single numbers often leads to errors in judgment. Fuzzy linguistic and triangular fuzzy numbers are used to rate factors in this paper, shown in Table 1 and Figure 1. Specifically, for the failure modes, higher occurrence and severity, higher score. For the rating of subsystems, assume that the best state of the art and environmental condition and the lowest intricacy of subsystems under the existing circumstance get the full marks (10), the closer to the limitations, the higher scores the subsystem rated.

To obtain the clear decision-numbers, the fuzzy rating results must be defuzzified. The current defuzzification methods mainly include the mean of maxima (MOM), center of area (COA) and α-cut etc. [18]. Different methods lead to various results. COA method is applied while considering the demand of simple practicality in allocation process, the expression of COA method is given as:

Table 1 Fuzzy ratio scale and membership function of linguistic terms

Linguistic variable	Triangular fuzzy number
Very low(VL)	(0,0,1)
Low(L)	(0,1,3)
Medium low(ML)	(1,3,5)
Medium(M)	(3,5,7)
Medium high(MH)	(5,7,9)
High(VH)	(7,9,10)
Very high(VH)	(9,10,10)

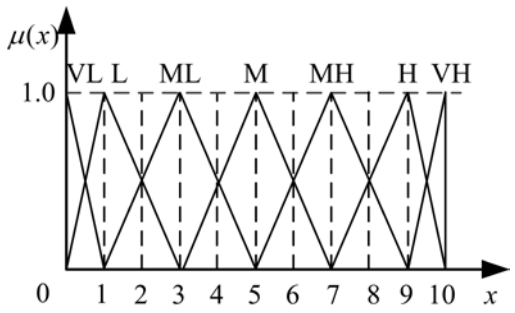


Fig. 1. Membership functions of triangular fuzzy numbers

$$x(a) = a_1 + \frac{1}{3}[(a_3 - a_1) + (a_2 - a_1)], \tag{22}$$

where *x(a)* is the defuzzified value, a_1, a_2, a_3 is the upper limit, most probable value and the lower limit value respectively.

Step 3 PR_i determination

Different severities of failure modes have different effects on system. To solve the unreasonable of factor weights in the allocation of RPN-based methods, the severity is modified as Eq.(23) based on the approach proposed by Kim et al. [15]. Not only does this method make up for the equal weight of factors, but solves the linearity of various severities which are criticized in RPN-based methods:

$$S'_{ij} = a^{S_{ij}}, \ a > 1. \tag{23}$$

where *a* is the risk coefficient related to the type of products, the more serious the failure effects of this product, the higher value of *a* must be selected.

The failure mode numbers, severity and occurrence of each failure modes in a subsystem codetermine the potential risk, where the single loss is depend on the severity of failure mode and the loss frequency in a certain time is determined by both the number of failure modes and occurrence. Therefore, we proposed that the potential risk of subsystems should expressed as:

$$PR_i = \sum_{j=1}^{N_i} O_{ij} S'_{ij} \cdot \tag{24}$$

Step 4 *C* determination

Plenty of papers show that it is not the simple linear relationship between costs and the improvement of systems reliability. Costs increase with the improvement of system reliability and would be a very high value while the reliability closes to the ultimate value under the current circumstance. Based on this property, the corresponding costs of system for its reliability we proposed is given as:

$$C'_i = \log_b(1 - \frac{C_i}{C_{\max}}), \tag{25}$$

$$C_i = SA_i \times SI_i \times EC_i, \tag{26}$$

where C'_i is the final cost rating of subsystem i. C_i is the defuzzification rating of subsystem i, the higher C_i indicates that the subsystem owns the higher reliability level and the lower potential for reliability improvement. $C_{\max}$ is the ultimate value of subsystem i in the current technological level. b is the cost coefficient where $b \in (0,1)$.

As shown in Figure 2, the cost described by Eq. (25) has following two features. Firstly, the higher reliability of system, the more costs it needed for the improvement of equal reliability ΔC, that is $\Delta C'_1 > \Delta C'_2$. Secondly, at the same level of reliability, different types of products have different effort coefficients, and the cost of raising the same reliability is different, that is $C'_{i2} > C'_{i3}$.

To avoid ignoring the smaller value of factor due to the larger value of others in allocation process, the potential risk and manufacturing costs of subsystems should kept in the same magnitudes, that is to say, PR_i and C'_i should satisfy the limitation of $10^{-1} \le PR_i/C'_i \le 10$ while b is limited as:

$$\frac{10\ln(1 - C_{i\min}/C_{\max})}{\exp(PR_{\min})} \le b \le \frac{\ln(1 - C_{i\max}/C_{\max})}{10\exp(PR_{\max})}, \tag{27}$$

Step 5 System allocation methods

The reliability of system is allocated to its components and is satisfied by the reliability combination of subsystems in the end. The basic target of reliability allocation is minimizing the possibility damages of system by a reasonable method which requires that the potential risks and the necessary reliability costs of subsystems must be weighted. The larger value of *PR*, the more serious the possible failure damage of a subsystem is. The smaller value of *C'*, the higher the potential for reliability improvement of a subsystem is. The lower failure rate must be assigned to the subsystem which has higher potential risk and lower manufacturing costs for the sake of optimal results:

$$n_i = \sum_{i=1}^{k} PR_i - PR_i + C'_i, \tag{28}$$

Plugging this into Eq.(3) , the final allocation weight is expressed by:

$$\omega_i = \frac{\sum_{i=1}^{k} PR_i - PR_i + C'_i}{\sum_{i=1}^{k}(\sum_{i=1}^{k} PR_i - PR_i + C'_i)}, \; i=1,2,\ldots,k, \tag{29}$$

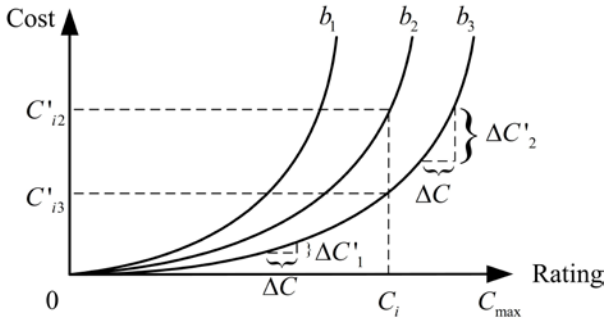


Fig. 2. Transformed cost rating

4. Illustrative example

To illustrate the effectiveness of proposed method further, the reliability allocation of spindle system of numerical control machine is employed in this section.

The spindle system is made up of spindle, bearing group, cooling system, broaching mechanism and rotation driving. Assume that the target failure rate of spindle system λ^*=0.002. Now three experts (E1, E2 and E3) are asked to rate the failure modes and subsystems by the linguistic variables as shown in Table 1. The rating results are shown in Table 2. The defuzzified results which are translated from linguistic variables to triangular numbers are expressed in Table 3.

Table 4 shows the allocations results of three methods. The results of RPN-based allocation method are calculated by Eq. (11) due to the attention of allocator is minimizing the potential risk of system. Results of proposed method are obtained at the circumstance of $a = \sqrt{e}$, b=0.997. The traditional allocation results are modified based on Eq. (5) for the reason that the rating principle of traditional method is opposite to this paper.

As shown in Table 4 and Figure 3, subsystems are allocated diverse failure rates under different methods. Cooling system is given the highest failure rate in the proposed method which is the same result of RPN-based approach. As shown in Table 3 and Table 4, cooling system has the highest rating of C_i which means owning the highest relative reliability, and it must cost more than others to increase the same reliability. Meanwhile, it has the lowest rating of PR_i which means having the minimal effects on system when failure happens. Therefore, it is more reasonable to assign the highest failure rate to cooling system than others.

The broaching mechanism is given the lowest failure rate in proposed method while the lowest failure is given to the cooling system and spindle in traditional and RPN-based method respectively. It can be explained that, the traditional allocation method is concentrates on the current reliability of system and assigns the lower failure rate to the subsystem with higher reliability, while RPN-based method focuses on the failure effects of subsystems to system and assigns the lower failure rate to the subsystem with more serious effects for minimizing the probable losses. Table 3 and Table 4 show that, cooling system has the highest rating of C_i and the spindle owns the highest mean value of failure modes rating which means that the cooling system has the highest relative reliability and the spindle has the most serious failure effect to system. Therefore, the lowest failure rating is given to cooling system and spindle respectively. However, both traditional and RPN-based approaches are considered only unilaterally, without optimizing allocation results. Though the spindle owns the largest *PR*, the necessary reliability costs are massive due to its high relative reliability, and it is more unreasonable to allocate the lowest failure rate to cooling system. The broaching mechanism owns a low relative reliability in subsystems which means having a big room for reliability improvement, it more necessary to pay more attention to it while *PR* is the second in subsystems. Therefore, it is the optimization

Table 2. Ratings of subsystems and failure modes assessed by experts

i	Subsystems	SA_i			EC_i			SI_i			Failure modes	O_{ij}			S_{ij}		
		E1	E2	E3	E1	E2	E3	E1	E2	E3		E1	E2	E3	E1	E2	E3
1	Spindle	MH	H	H	VH	H	MH	MH	VH	H	Orientation error (FM_{11})	H	MH	H	M	M	MH
											Accuracy error (FM_{12})	ML	L	L	M	MH	M
											Abnormal sound (FM_{13})	VH	H	VH	ML	M	M
											Over-heat (FM_{14})	ML	L	L	L	VL	L
2	Bearing group	VH	H	VH	H	M	MH	H	VH	H	Excessive clearances (FM_{21})	H	M	L	H	ML	M
											Ball drops out (FM_{22})	VL	VL	VL	VH	VH	VH
3	Cooling system	H	H	H	VH	H	VH	VH	H	H	Cannot refrigerate (FM_{31})	M	L	ML	MH	ML	ML
											Leak (FM_{32})	VH	VH	H	L	L	ML
4	Broaching mechanism	H	MH	MH	H	MH	MH	H	H	MH	Loose (FM_{41})	MH	H	ML	M	L	MH
											Fracture (FM_{42})	L	VL	L	H	VH	VH
5	Rotation driving	MH	H	MH	H	H	H	H	H	MH	Jam (FM_{51})	L	M	ML	MH	MH	H

Table 3. Defuzzified fuzzy ratings of subsystems and failure modes

i	SA_i	EC_i	SI_i	C_i	Failure modes	O_{ij}	S_{ij}	$\sum_{j=1}^{k} O_i S_i$
1	8.11	8. 67	8. 67	609.6198	FM_{11}	8.34	5.67	98.6734
					FM_{12}	1.79	5.67	
					FM_{13}	9.11	4.33	
					FM_{14}	1.79	1.00	
2	9.33	7.00	9.00	587.7900	FM_{21}	5.00	5.56	30.9911
					FM_{22}	0.33	9.67	
3	8.67	9.33	9.00	728.0199	FM_{31}	3.00	4.33	30.2079
					FM_{32}	9.11	1.89	
4	7.56	7.56	8.11	463.5157	FM_{41}	6.22	4.44	36.7268
					FM_{42}	1.00	9.11	
5	7.56	8. 67	8.11	531.5716	FM_{51}	3.00	7.67	23.01

result to allocating the minimum failure rate to broaching mechanism.

Table 5 and Figure 4 show the allocated failure rate of subsystems under various cost coefficient b. In order to satisfy the demand that PR_i and C_i should kept in the same magnitudes, b is limited to the interval of [0.9867,0.9998] by using Eq.(27) . As shown in Table 5 and Figure 4, the allocated failure rates change with different value of b. When b is close to the lower limit (b=0.990), the lowest failure rate is assigned to spindle while the highest is assigned to cooling system. When b is close to the upper limit (b=0.999), the lowest failure rate assigned to broaching mechanism while the assigned failure rate of spindle rises to the third. This can be explained that b must get a higher value in Eq. (25) while the production costs are higher or the designers are more concerned about costs than the failure effects on system, in other word, b must get a lower value in Eq. (25) while the failure effects are more serious than its manufacturing costs. When b=0.990, the necessary reliability costs are considered lesser than the losses of failures by designers, failure effects of

Table 4. Comparison of the results obtained from different allocation methods

i	PR_i	$\sum_{i=1}^{k} PR_i$ -PR_i	C'_i	ω_i	λ^*_i		
					Traditional	RPN-based	Proposed
1	254.8591896	462.9606091	313.0741438	0.177486987	0.000395632	**0.000372372**	0.000354974
2	122.1218372	595.6979614	294.9641172	0.203703414	0.000399369	0.000419830	0.000407407
3	49.58213563	668.2376630	433.3574532	0.251945931	**0.000375361**	**0.000421856**	**0.000503892**
4	152.3763532	565.4434454	207.2611430	0.176725345	**0.000420645**	0.000404992	**0.000353451**
5	138.8802830	578.9395156	252.4112766	0.190138324	0.000408994	0.000380951	0.000380277
Total	717.8197986	2871.279195	1501.068134	1		0.002	

Table 5. Influence of cost coefficient b on the allocation results

i	λ^*_i			
	b=0.990	b=0.995	b=0.997	b=0.999
1	**0.000335271**	**0.000345062**	0.000354974	0.000380302
2	0.000411972	0.000409703	0.000407407	0.000401539
3	**0.000480593**	**0.000492170**	**0.000503892**	**0.000533843**
4	0.000377952	0.000365777	**0.000353451**	**0.000321954**
5	0.000394213	0.000387288	0.000380277	0.000362362
Total	0.002			

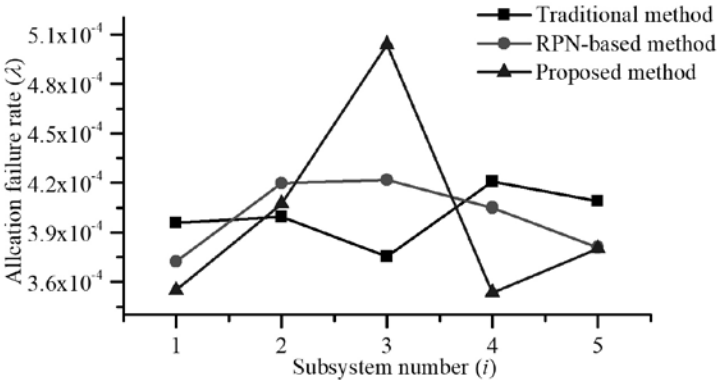


Fig. 3. Comparison of the results obtained from different allocation methods

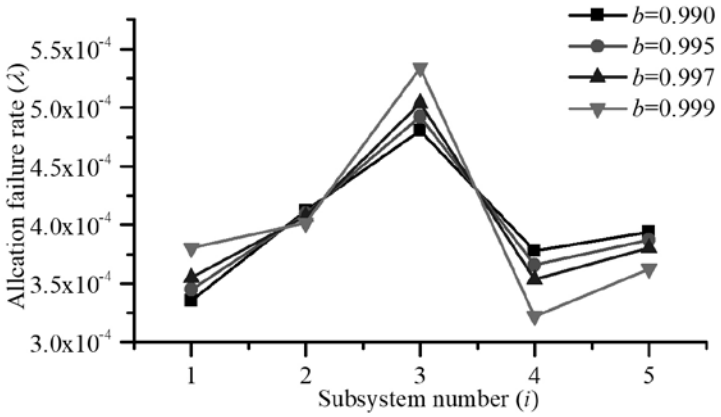


Fig. 4. Influence of cost coefficient b on the allocation results

subsystems is predominant during the allocation process so that the rank of allocation results is similar to the results of RPN-based. As for the difference between broaching mechanism and rotation driving can be explain that the values in RPN-based method are mean values, and the weight of broaching mechanism is raised after averaged. When b=0.999, the designers are more focused on design costs. The failure effects of subsystems are slighter relatively while the costs of subsystems are predominant. The ranks of allocation results are opposite to traditional method. When b exceed the limitation, the lower weight will ignored due to the far less than the higher, which would decrease the credibility of the allocation results.

5. Conclusions

This paper provided a comprehensive reliability allocation method considering failure effects and the necessary costs of system in a specific reliability. The potential risks and reliability costs of subsystems are considered as the allocation factors, the modified RPNs are used to represent the potential risks of subsystems, and the reliability costs model is created by using relative reliability of subsystems. An allocation model is constructed for the purpose of optimizing results which solves the weaknesses of incomplete considerations and poor practicability in the existing reliability allocation methods. It is more flexible while the risk and cost coefficients are considered which can be adjusted with various purposes or allocating objects. Uncertainty factors in allocation process are accounted by fuzzy method and the presented value range of risk and cost coefficients ensures the balance of weight factors, which both enhance the credibility of results.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (No. 51575070), the National Major Scientific and Technological Special Project for "High-grade CNC and Basic Manufacturing Equipment" of China (2015ZX04003-003, 2016ZX04004-005), the Fundamental Research Funds for Central Universities (No.106112017CDJXY110006).

References

1. Bracha VJ. The methods of reliability engineering .Machine Design 1964; 7:70–6.
2. Certa A, Hopps F, Inghilleri R, et al. A Dempster-Shafer Theory-based approach to the Failure Mode, Effects and Criticality Analysis (FMECA) under epistemic uncertainty: application to the propulsion system of a fishing vessel. Reliability Engineering & System Safety 2017; 159(69):79, https://doi.org/10.1016/j.ress.2016.10.018 .
3. Chen J, Duan M, Zhang Y. Decision-making of spare subsea trees with multi-restrictive factors in deepwater development. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18 (4): 590–598, http://dx.doi.org/10.17531/ein.2016.4.14.
4. Dale C J, Winterbottom A. Optimal Allocation of Effort to Improve System Reliability. IEEE Transactions on Reliability 1986; 35(2):188-191, https://doi.org/10.1109/tr.1986.4335401 .
5. Department of the Army. TM 5-689-4. Failure modes, effects and criticality analysis (FMECA) for command, control, communications, computer, intelligence, surveillance, and reconnaissance (C4ISR) facilities 2006 Sep.
6. Elegbede A O C, Chu C, Adjallah K H, et al. Reliability allocation through cost minimization. IEEE Transactions on Reliability 2003; 52(1):106-111, https://doi.org/10.1109/tr.2002.807242 .
7. Gong Q X. Reliability Engineering Handbook. National Defense Industry Press, 2007.

8. Gölbaşı O. Risk-Based Reliability Allocation Methodology to Set a Maintenance Priority Among System Components: A Case Study in Mining[J]. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19(2):191-202, http://dx.doi.org/10.17531/ein.2017.2.6.
9. Heydorn R P. Reliability Engineering Handbook. Prentice-Hall, 2001.
10. Huang H Z, Liu Z J, Li Y, et al. A warranty cost model with intermittent and heterogeneous usage. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2008; 40(4):9-15.
11. Itabashi-Campbell R R, Yadav O P. System Reliability Allocation based on FMEA Criticality// SAE World Congress & Exhibition 2009.
12. Jaśkowski P. Methodology for enhancing reliability of predictive project schedules in construction. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17(3):470-479, https://doi.org/ 10.17531/ ein.2015.3.20 .
13. Karmiol ED. Reliability apportionment. Preliminary Report EIAM 5, Task II. General electric. Schenectady, NY 1965.
14. Kececioglu D. Reliability engineering handbook (vol. 1). Prentice Hall, 1992.
15. Kim K O, Yang Y, Zuo M J. A new reliability allocation weight for reducing the occurrence of severe failure effects. Reliability Engineering & System Safety 2013; 117(117):81–88, https://doi.org/10.1016/ j.ress.2013.04.002 .
16. Kuo W, Prasad V R, Tillman F A, et al. Optimal Reliability Design: Fundamentals and Applications. Microelectronics Journal 2001; 32(10):911-911.
17. Li J, Guo J Z, Zhou H Q, et al. Research on the Method of Reliability Allocation of Diesel Engine Base on the Cost Function. Applied Mechanics & Materials 2012; 271-272(2446):1115-1120, https://doi.org/10.4028/www.scientific.net/amm.271-272.1115 .
18. Liu H C, You J X, You X Y, et al. A novel approach for failure mode and effects analysis using combination weighting and fuzzy VIKOR method. Applied Soft Computing 2015; 28(C):579-588, https://doi.org/10.1016/j.asoc.2014.11.036 .
19. Liu Y, Yu W, Li Y, et al. Reliability allocation based on interval analysis and grey system theory. China Mechanical Engineering 2015; 26(11):1521-1526, https://doi.org/10.1016/j.infsof.2016.09.010 .
20. O'Hagan M. Aggregating Template or Rule Antecedents In Real-time Expert Systems With Fuzzy Set Logic// Asilomar Conference on. IEEE Xplore 1988; 681-689, https://doi.org/10.1109/ acssc.1988.754637 .
21. Qiu X, Ali S, Yue T, et al. Reliability-Redundancy-Location Allocation with Maximum Reliability and Minimum Cost Using Search Techniques. Information & Software Technology 2016; 82:36-54, https://doi.org/10.1016/j.infsof.2016.09.010 .
22. Todinov M T. Risk-based reliability allocation and topological optimization based on minimizing the total cost. International Journal of Reliability & Safety 2007; 1(4):489-512(24), https://doi.org/10.1504/ijrs.2007.016261 .
23. Wang Y, Yam R C M, Zuo M J, et al. A comprehensive reliability allocation method for design of CNC lathes. Reliability Engineering and System Safety 2001; 72(3):247-252, https://doi.org/10.1016/ s0951-8320(01)00018-7 .
24. Xiao N, Huang H Z, Li Y, et al. Multiple failure modes analysis and weighted risk priority number evaluation in FMEA. Engineering Failure Analysis 2011; 18(4):1162-1170, https://doi.org/10.1016/ j.engfailanal.2011.02.004 .
25. Yadav O P, Zhuang X. A practical reliability allocation method considering modified criticality factors. Reliability Engineering & System Safety 2014; 129:57-65, https://doi.org/10.1016/j.ress.2014.04.003 .
26. Yadav O P. System reliability allocation methodology based on three- dimensional analyses. International Journal of Reliability & Safety 2007; 1: 360–75, https://doi.org/10.1504/ijrs.2007.014969 .
27. Zhang G B, Jian L, Wang G Q. Fuzzy reliability allocation of CNC machine tools based on task. Computer Integrated Manufacturing Systems Cims 2012; 18(4):768-774.

Hui YU
Genbao ZHANG
Yan RAN
College of Mechanical Engineering and State Key Lab Mech Transmiss, Chongqing University, Chongqing 400044, China

Mengqi LI
College of Mechanical and Energy Engineering, Shaoyang University, Shaoyang 422000, China

Yang WANG
College of Mechanical Engineering and State Key Lab Mech Transmiss, Chongqing University, Chongqing 400044, China

E-mails: vander_yu@163.com, gen.bao.zhang@263.net, ranyan@cqu.edu.cn, sciencefield@126.com, 175473324@qq.com

Article citation info:
FERDYNUS M, KOTEŁKO M, KRAL J. Energy absorption capability numerical analysis of thin-walled prismatic tubes with corner dents under axial impact. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 252–259, http://dx.doi.org/10.17531/ein.2018.2.10.

Mirosław FERDYNUS
Maria KOTEŁKO
Jan KRAL

ENERGY ABSORPTION CAPABILITY NUMERICAL ANALYSIS OF THIN-WALLED PRISMATIC TUBES WITH CORNER DENTS UNDER AXIAL IMPACT

NUMERYCZNA ANALIZA ENERGOCHŁONNOŚCI CIENKOŚCIENNYCH SŁUPÓW PRYZMATYCZNYCH Z PRZETŁOCZENIAMI

The paper presents results of a parametric study into energy absorption capability of thin-walled square section columns with redrawn dents, subjected to axial impact compressive load. Thin-walled aluminum tubes with four dents in the corners were under investigation. The varying parameters were the dent's depth and distance of the dent to the base. The study was performed using Finite Element numerical code. Three crashworthiness indicators were examined: peak crushing force, crash load efficiency and stroke efficiency. The numerical results are shown in load-shortening diagrams, as well as diagrams and maps of crashworthiness indicators. It was found, that the main factor influencing a crushing mode and, subsequently, energy absorption capability, is a dent depth. The dent distance from the base is of less importance. Also a position of a dent, either at the bottom, or at the top base (the load application point) does not influence the crushing behavior significantly. For the deepest dents the relative increase of crash load efficiency (CLE) amounts 25% in comparison with the column without dents.

Keywords**:** *crashworthiness indicators, energy absorber, thin walled member.*

W artykule przedstawiono wyniki badań numerycznych zdolności pochłaniania energii energoabsorberów w postaci cienkościennych słupów o przekroju kwadratowym z wgłębieniami, poddanych osiowym obciążeniom udarowym. Badano wpływ parametrów geometrycznych oraz położenia inicjatorów zgniotu w postaci walcowych przetłoczeń w narożach na zachowanie się konstrukcji oraz właściwości energoabsorbcyjne (współczynnik efektywności zgniotu- Ste oraz procentowy stosunek siły średniej do maksymalnej - CLE). Obliczenia numeryczne prowadzono z wykorzystaniem MES, programu Abaqus 6.14. Wyniki przedstawiono w postaci charakterystyk obciążenie – skrócenie oraz diagramów i wykresów. Stwierdzono, że istotny wpływ na zachowanie się konstrukcji podczas uderzenia oraz jej energochłonność ma głębokość przetłoczenia, mniej istotne jest jego położenie. W przypadku słupów z najgłębszymi przetłoczeniami względny wzrost współczynnika CLE, w porównaniu z wynikami uzyskanymi dla słupa gładkiego wynosi 25%.

Słowa kluczowe*: wskaźniki energoabsorbcyjne, absorber energii, konstrukcje cienkościenne.*

1. Introduction

Increasing number of impacting events of many types like traffic accidents, collisions of ships or collisions of a ship either with an iceberg or ship grounding on a narrow rock, etc. induced the rapid development of the impact crashworthiness dealing with research into impact engineering problems, particularly in the field of dynamic response of structures in the plastic range and the design of energy absorbers. Since demands of general public of the safe design of components of vehicles, ships, etc. have increased substantially in the last few decades, a new challenge appeared to design special structural members which would dissipate the impact energy in order to limit the deceleration and finally to stop a moveable mass (e.g. vehicle) in a controlled manner. Such a structural member termed the energy absorber converts totally or partially the kinetic energy into another form of energy. One of the possible design solutions is the conversion of the kinetic energy of impact into the energy of plastic deformation of a thin-walled metallic structural member.

In the early sixties of the 20th century, automotive safety regulations stimulated the development of the new concept of a crashworthy (safe) vehicle that had to fulfil the integrity and impact energy management requirements [2]. A designer of any impact attenuation device must meet two main, sometimes contrary requirements: The initial collapse load has to be not too high in order to avoid unacceptably high impact deceleration of the vehicle. On the other extreme, the main requirement is a possibly highest energy dissipation capacity, which may not be achieved if the collapse load of the impact device is too low. The latter may result in dangerously high occupant "ridedown" decelerations.

Thus, maximizing energy absorption and minimizing peak to mean force ratio by seeking for the optimal design of these components are of great significance.

There are numerous types of energy absorbers of that kind that are cited in the literature [4,6]. Namely, there are steel drums, thin tubes or multi-corner columns subject to compression, compressed frusta (truncated circular cones), simple struts under compression, sandwich plates or beams (particularly honeycomb cells) and many others. Among all those design solutions, mentioned above, thin-walled metal tubes are widely used as energy absorption systems in automotive industry due to their high energy absorption capability, easy to fabricate, relatively low price and sustainability at collapse.

In the case of thin-walled members subjected to axial compression, which act as energy absorbers, a substantial issue is such a design, which promotes a progressive buckling mechanism, stimulating the highest energy absorption capacity. One of the possible design solutions is the application of a trigger (notch or dent), which releases the most desirable crushing mechanism. There are numerous published results of research concerning energy absorption of thin-walled tubes [2,4,6], however, few deal with tubular structures with dents or other flaws.

2. State of the art: an overview

As was mentioned above, metal tubes, particularly of prismatic cross-section, subjected to impact load, have been investigated since 60-ties of the 20th century [3]. Many efforts have been made so far by several researchers to reduce the peak load at the initial stage of the crushing process and simultaneously to increase energy absorption capacity in such members, acting as energy absorbers [2,6].

At the turn of centuries, a new class of tubular energy absorbers have been introduced, namely axially loaded thin-walled foam-filled metal columns made of different foam material, like polyurethane, metal, polypropylene, etc.. Energy absorbed during the crushing process by such members consists of the energy dissipated in the metal tube and the energy dissipated in the filler.

The problem of the crushing behavior of such members was initially investigated by Hanssen et al. [11,12]. Then, Wang et al. [20] published the results concerning closed-hat section columns filled with aluminum foam. They derived a theoretical model of such columns and developed a constitutive material model of the aluminum foam on the basis of experimental tests.

In last two decades numerous research results concerning the problem of foam-filled metal columns acting as energy absorbers have been published [6]. Because of many parameters influencing energy absorption capability and more and more complex structural design of such members, optimization procedures occurred to be necessary. One of the first attempts in this area were published by Zarei and Kroger [22], who performed the optimization of the foam-filled aluminum tube.

In the first decade of 21th century, multi-cell columns have been introduced as energy absorbers, hollow and foam-filled. Chen and Wierzbicki [7] published theoretical results concerning crushing behavior of such structural members. They investigated the axial crushing of multi-cell hollow columns, using analytical and numerical methods and of double-cell and triple-cell foam-filled columns, using FE method. Very recently, Yin et al. [21] performed multi-objective crashworthiness optimization of foam-filled multi-cell columns using different metamodel techniques. The optimization criteria were maximum energy absorption and minimum peak crushing force. They carried out the multi-objective optimization using four metamodels and stated, that the most effective one was Radial Basis Function (RBF) model.

In last few years, research was carried out into another solution, namely the application of multi-cornered thin-walled tubes as energy absorbing members [6]. Abbasi et al. [1] investigated the crushing behavior and energy absorption of hexagonal, octagonal and 12-edge thin-walled columns. Reddy et al. [18] continued to analyze the similar problem for 12-edge columns of different geometry. Ali et al. [5] analyzed the axial crushing of pentagonal and cross-shape thin-walled tubes. Another solution, bitubal energy absorber was investigated experimentally under quasi-static load by Sharifi et al. [19].

Very recently, W. Liu et al. [17] performed the crashworthiness multi-objective optimization of sandwich star-shaped tubes. They investigated a multi-wall structure, with a star-shaped tube between two circulars.

Another alternative solution are flaws or dents acting as triggers. A trigger may induce the most desirable crushing (collapse) mode, leading to higher energy absorption and mean to peak crushingforce ratio.

Very few published papers deal with tubular structures with dents, dimples or other flaws. An interesting solution has been presented by Z. Yang [23]. The subject of investigation was the crushing behavior of a thin-walled circular tube with internal gradient grooves. The authors fabricated stainless steel thin-walled tube with preset internal circumferential rectangular groove defects using SLM 3D printing method. They observed double buckling-splitting crushing mode. Empty and foam-filled circumferentially grooved thick-walled circular tubes under axial low velocity impact were investigated theoretically and experimentally by A. Darvizeh et al. [8]. In [9] the effect of discontinuity size on the energy absorption performance of square profiles was reported. In [16] the authors studied the behavior of tubular columns with dents, but these flaws were treated as imperfections coming from damage.

One of the first attempts to apply columns with dents as energy absorbers, namely a static analysis of axially compressed square section tube with dents in the corners is presented in [10]. The subject of the investigation was a square-section column with periodically (along the height) situated dimples (redrawn dents) in the corners. FE simulations of crushing behavior were carried out. Very recently, Yang et al [15] investigated circular tubes with periodically situated (both along the circumference and the height) ellipsoidal dimples, subjected to axial crushing force. FE simulations of crushing behavior were performed and experimentally validated by quasi-static tests on 3D printed brass tubes.

The above state of the art overview shows that no research has been undertaken so far into crushing behavior and crashworthiness of thin-walled columns with a prismatic cross-section and non-periodically situated dents. In this paper the crushing behavior of such structural members subjected to axial impact load is investigated.

3. Subject and aim of the investigation

3.1. Structural geometry

The subject of this investigation was a thin-walled square section aluminum tube with four indentations in the corners, formed by means of cylindrical redrawing. The tubes with the dimensions □70x2 and height l=335 mm were investigated.

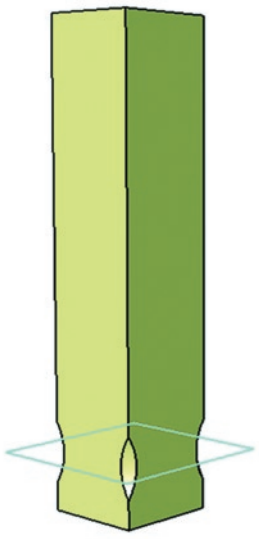


Fig. 1. Column model with dents generated in Catia v5 system

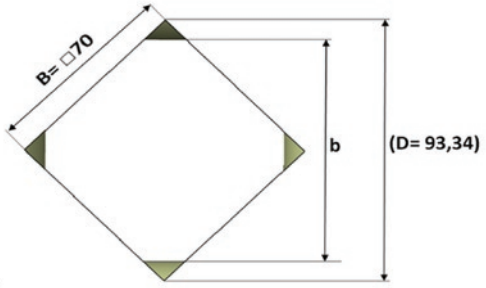


Fig. 2. Cross-section of the column in the mid-surface of the dent (D=93,34 mm)

The process of modeling of the column with dents was realized with the Catia v. 5 software package in the Generative Shape Design module. The geometrical model was exported to the Abaqus code module. The dent's geometry was characterized by the main radius R=50 mm and relative dent depth with respect to the dimension **b** of the main diagonal (Fig.2) in the column's cross-section (from 5% up to 30%). The dents were made at the bottom of the column (Fig.1,3) or at the top (Fig.3). The models with the dents at the bottom of the column were designated by the symbols from A05_X to A30_X, where the first number stands for the relative depth of the dent (in percentage) while X stands for the distance of the dent from the bottom (Fig.4). Similarly, the models with dents situated at the top of the column (where the impact load was applied) were designated by the column symbols C05_X to C30_X, where X was measured from the top of the column. The column with smooth walls (without dents) was denoted as SM.

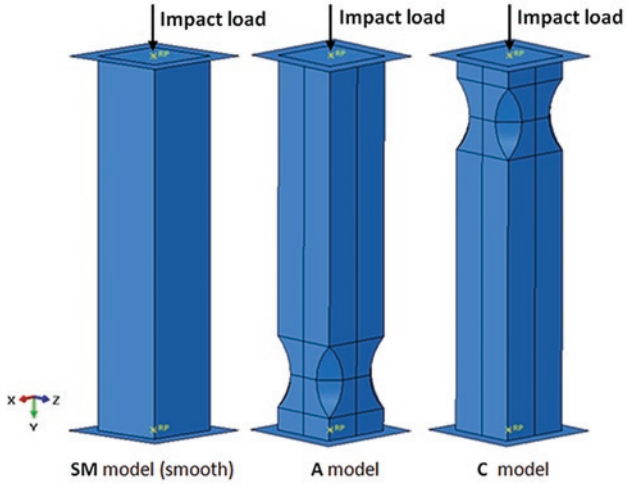


Fig. 3. Investigated column models

A05_X	A10_X	A15_X	A20_X	A25_X	A30_X
b=88.67	b=84.01	b=79.34	b=74.67	b=70.01	b=65.34
X= 5; 10; 15; 20; 25; 30					

Fig. 4. Geometric parameters of A models

4. FE model

Crushing behavior analysis of the investigated columns was performed on the basis of FE simulations. The FE explicit analysis was carried out using Abaqus 6.13 code. The column was situated between two rigid plates (Fig. 3). An FE model of the column was created using 4-node shell elements S4R and it was subjected to impact load of the kinetic energy E= 7.5 kJ, which corresponds to the mass m= 150 kg dropping with the initial velocity V_0= 10 m/s.

The columns were assumed to be made from the aluminum alloy EN AW6060- T6 (σ_Y=175 MPa, σ_{ult}=250 MPa, ν=0.33). Since aluminum alloys do not display a significant sensitivity to the strain rate, a bi-linear material model was applied, neglecting an influence of the strain rate, but taking into account the strain hardening (E=70000MPa, E_t=937,5MPa).

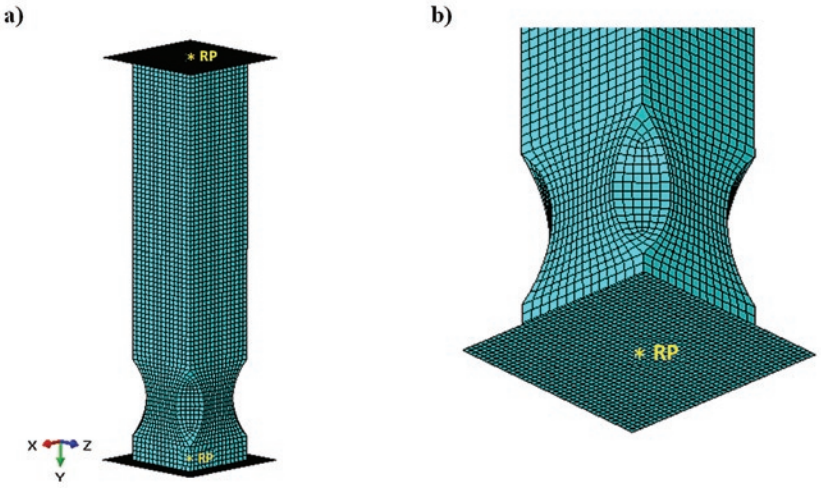


Fig. 5. FE model of the investigated column

In the rigid plates reference points RP were created, where the impact force (at the bottom point) and acceleration, velocity, displacement (at the top point) were recorded. The reference points are shown in Fig. 5. All degrees of freedom were taken (switched off) at the bottom rigid plate, while at the top plate only one degree of freedom, in the axial direction y, was possible.

5. Crushing behaviour

The results of the FE calculations reveal, that the failure mode of the investigated flawed columns depends mainly on the redrawing depth (dent depth). Three following main modes of failure were observed: For the smallest values of the dent depth (A05), regardless of the dent position (X), the crushing process was initiated by the local plastic mechanism (first local fold) situated at the center of the column (Fig.6). Also for A10, but for the position of dent X = 5-15, the same failure mode was observed. For medium values of dents depth (A10, A15) and X≤20 mm, the crushing process was triggered by the local plastic mechanism situated above the dent (Fig.7). For the largest values of dent depths (A20, A25, A30), regardless of the position of the dent, crushing was initiated by a local plastic mechanism situated exactly at the dent. These three types of failure are denoted below as mode I, mode II and mode III, respectively.

The specification of the crushing modes of columns A in terms of dent depth and position of the dent is shown in Table 1.

A preliminary FE study indicated, that the position of the dent, either at the bottom of the column (A) or at the top (C), does not influence the column's structural behavior significantly, which is shown in Fig. 9 and subsequent Fig. 10 and 11. The

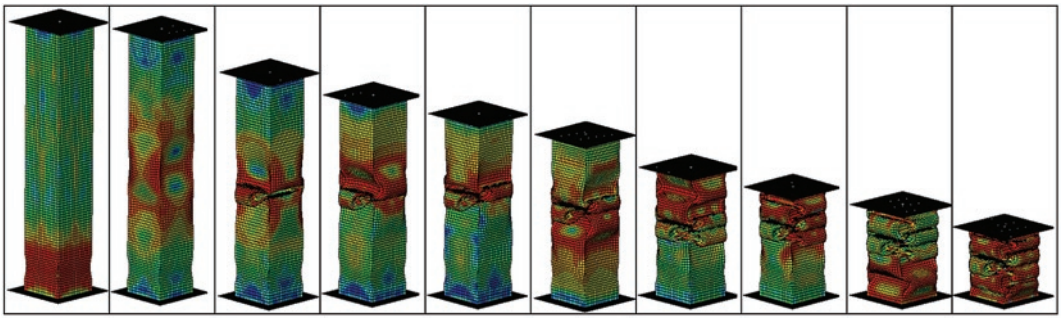


Fig 6. Crushing mode I (column A05_15)

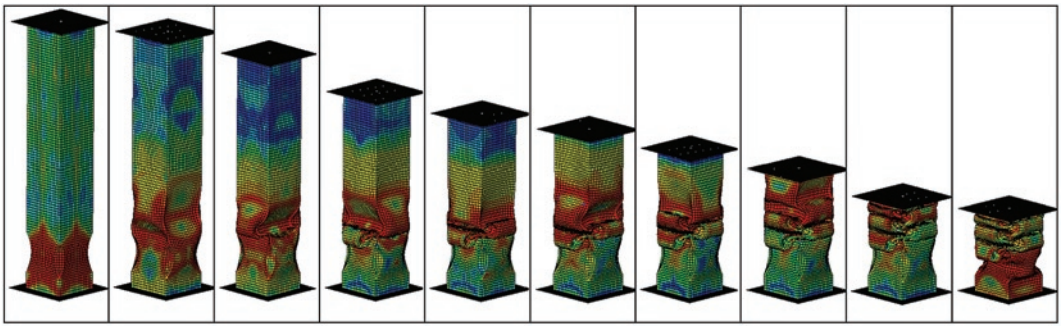


Fig.7. Crushing mode II (column A15_20)

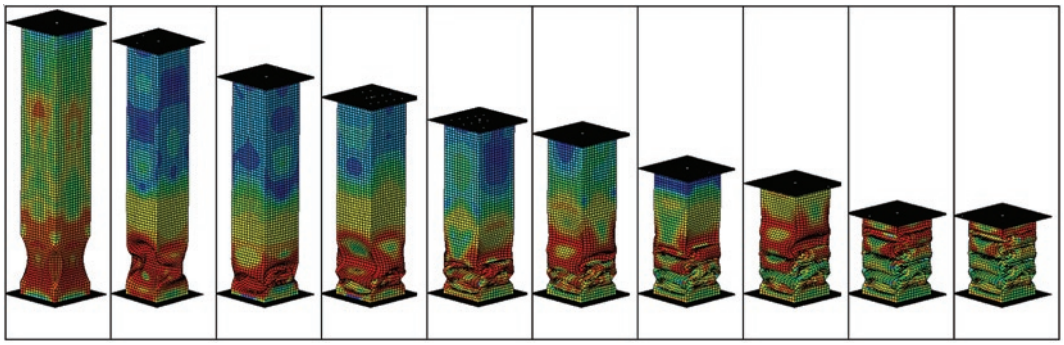


Fig. 8. Crushing mode III (column A20_10)

Table 1. Specification of columns A crushing modes

	Model					
	A05_X	**A10_X**	**A15_X**	**A20_X**	**A25_X**	**A30_X**
X=5	I	I	II	III	III	III
X=10	I	I	II	III	III	III
X=15	I	I	II	III	III	III
X=20	I	II	II	III	III	III
X=25	I	III	III	III	III	III
X=30	I	III	III	III	III	III

only more significant differences were observed with respect to the PCF factor (larger for the C option). Thus, a further detailed parametric study was focused on the column type A.

In all load-shortening diagrams the ordinate indicates the load value RF2 recorded at the lower reference point RP shown in Fig. 5. The abscissa indicates the vertical displacement U2 recorded at the upper reference point RP, also shown in Fig.5.

6. Parametric study

There are several crashworthiness indicators [13,14] used to evaluate the crashworthiness of the energy absorbing structure (energy absorber). In the present analysis, the following indicators are applied: energy absorption (EA), mean crushing force (MCF), crash load efficiency (CLE), stroke efficiency (St_e).

The typical crushing force – displacement curve for a thin-walled member subjected to axial impact is shown in Fig.12.

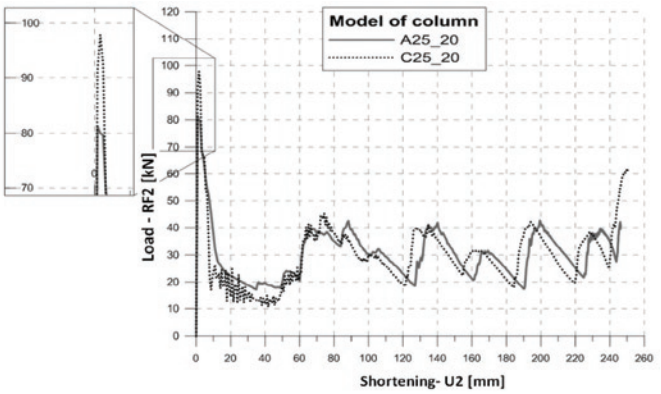


Fig. 9. Comparative load-shortening diagrams of column model A25_20 and C25_20

Table 2. Comparison of crashworthiness indicators for columns A and C

Model	U2max [mm]	PCF [kN]	MCF [kN]	St_e [-]	CLE [%]
SM	228,98	104,469	32,77	0,316	31,37
A25_20	246,27	87,557	30,48	0,265	34,79
C25_20	250,35	97,90	30,01	0,253	30,65

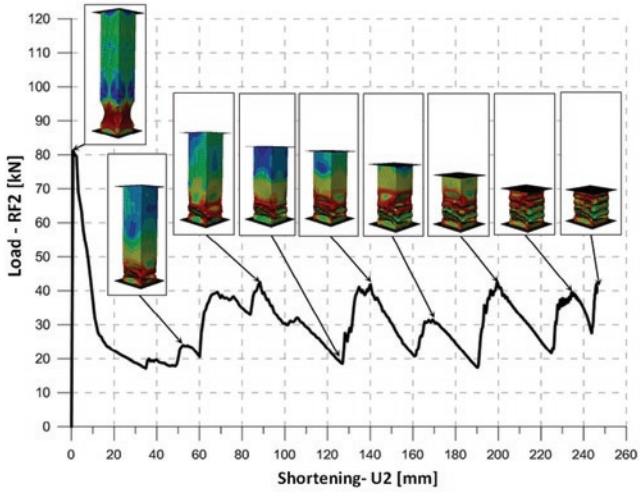


Fig 10. Load-shortening diagram of column model A25_20 (indication of characteristic failure stages)

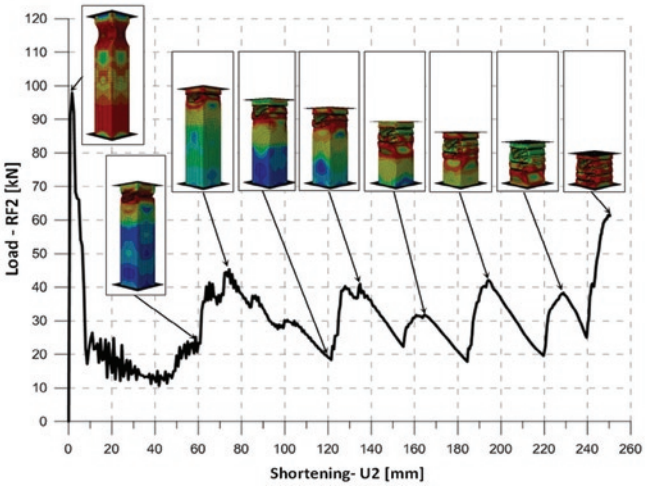


Fig. 11. Load-shortening diagram of column model C25_20 (indication of characteristic failure stages)

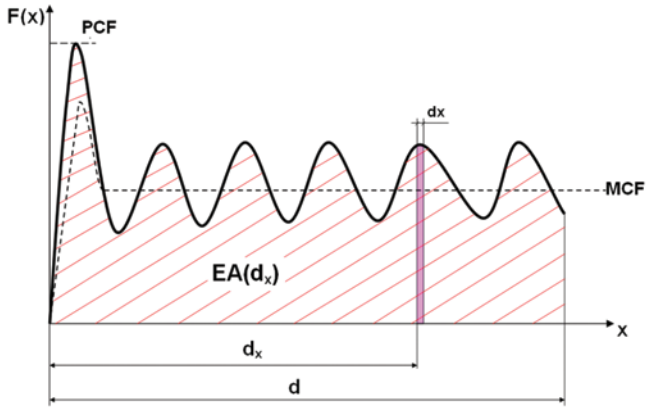


Fig. 12. Exemplary load-shortening diagram of thin-walled column under axial impact

$$EA(d_x) = \int_0^{d_x} F(x)dx \tag{1}$$

where d_x is a crushing distance (see Fig.12).

The mean crushing force (MCF) for a given crushing deformation d_x is calculated as:

$$MCF = \frac{EA(d_x)}{d_x} \tag{2}$$

The crash load efficiency is a mean crushing force (MCF) to peak crushing force (PCF – see Fig.12) ratio:

$$CLE = \frac{MCF}{PCF} \cdot 100\% \tag{3}$$

The stroke efficiency Ste is defined as follows:

$$St_e = \frac{L_o - U}{L_o} \tag{4}$$

where:

L_0 – is the initial length (characteristic dimension) of the member [mm],

U – is the maximum shortening (maximum characteristic deformation) of the member [mm]

A detailed parametric study into an optimal design of depth and position of the dents with respect to the column's energy absorption capability, namely crashworthiness indicators mentioned above (MCF, CLE and St_e), was performed on the basis of the FE simulations results.

As was mentioned above, the preliminary study showed, that the crushing behavior of columns A (bottom dent) and C (top dent) does not differ. Also, the values of crashworthiness indicators for these two types of columns are similar (except PCF), which is presented in Table 2. The difference in the Ste factor amounts to about 4.5%, while the CLE factor is larger about 12% for the A option. Since the A option (trigger at the bottom) is more desirable, the results presented below concern columns type A only.

Fig. 13. shows the load-shortening diagrams for columns A with different dent depths and the same dent position, which exhibited

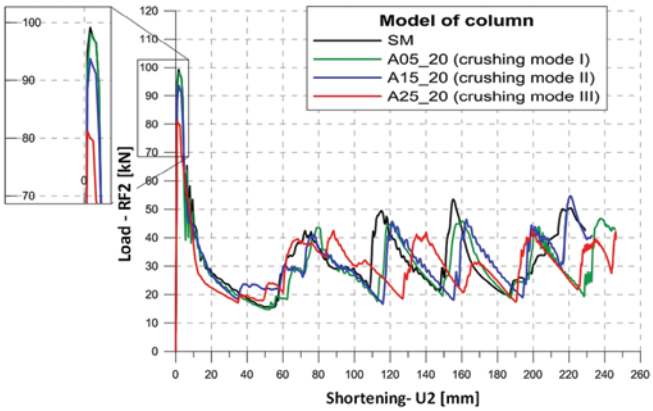


Fig. 13. Load-shortening diagrams for columns A exhibiting three different crushing modes (constant value of X=20 mm)

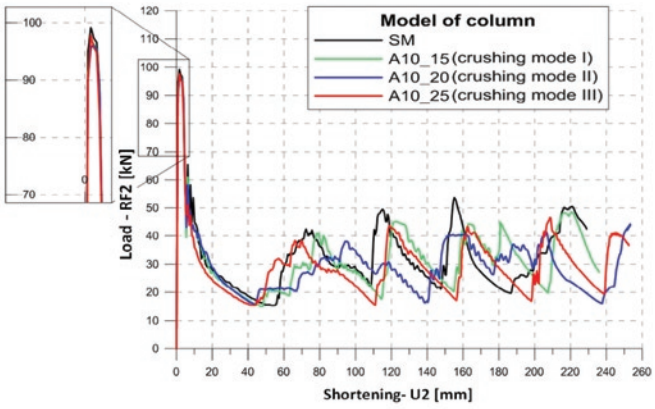


Fig. 14. Load-shortening diagrams for columns A exhibiting three different crushing modes (constant value of relative dent depth = 10%)

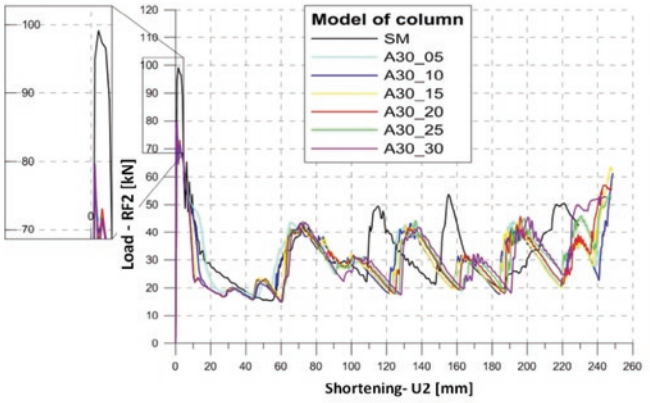


Fig. 15. Load-shortening diagrams for columns A exhibiting crushing mode III (constant value of relative dent depth = 30%)

three different crushing modes (see Table 1). The character of these diagrams is nearly the same. A significant decrease in the peak crushing force (PCF) is observed in comparison with the column without dents (SM). For the columns A10 (with the relative dent depth of 10%) also three crushing modes were observed, depending on the dent situtation X. Comparative load-shortening diagrams for three columns A10 of three X positions corresponding to three crushing modes are presented in Fig. 14. Also in that case the character of three diagrams is the same. In „blow-up" piece of the diagram one can notice, that PCF depends mainly on redrawing depth and, in minor extent, on X position.

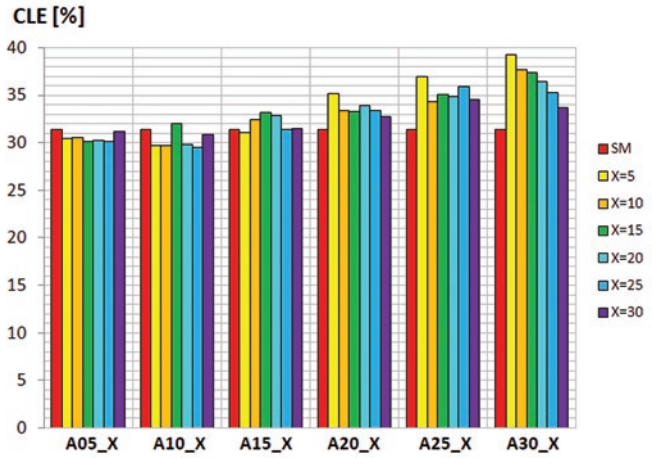


Fig. 16. Crash load efficiency in terms of relative dent depth

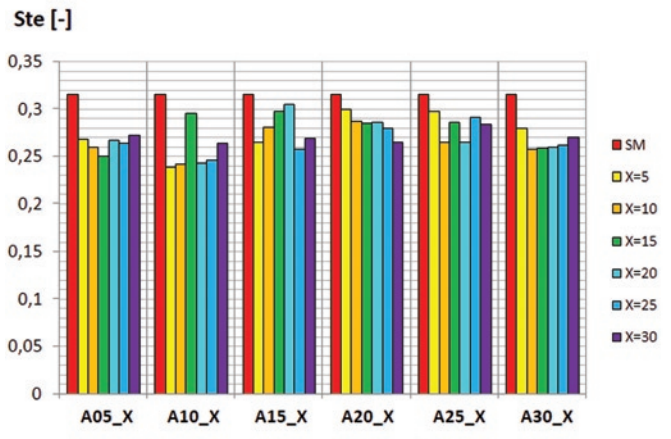


Fig. 17. Stroke efficiency in terms of relative dent depth

Fig. 15 shows a comparison of the load-shortening diagrams for columns A with the deepest dents (A30) and different X positions. These columns exhibit only the crushing mode III only.

The largest decrease of PCF was obtained for the columns with largest dent depth, which corresponds to crushing mode III and the smallest values of X (Table 3). The values of stroke efficiency (Ste) are shown in Table 4 and Fig. 17. This indicator depends on both the dent depth (A%) as well as on its position X. The smallest values are obtained for the relative dent depth (A10).

One of the most important crashworthiness indicators, both for the energy absorption capacity and for biomechanical reasons, is the CLE (mean to peak crushing force ratio). An ideal absorber should, if possible, have the CLE equal to 100%. For the columns under investigation the highest values of CLE were obtained for crushing mode III, corresponding to the deepest dents (A30) and for the smallest values of X, which is shown in Table 5 and Fig. 19. In Table 5, the shaded, grey areas visualize CLE values higher than those reperesenting the smooth column SM. The increase of the CLE in comparison with the column without dents (SM) in terms of the relative dent depth, for selected positions of X is shown in Fig. 16.

7. Conclusions

The paper reported the results of a parametric study into the energy absorption capability of thin-walled square section columns with dents. The varying parameters were the depth of the dent and the distance of the dent to the base. In addition to this, different positions of the dent,

Table 3. Peak crushing force

PCF [kN]	Column						
	SM	A05_X	A10_X	A15_X	A20_X	A25_X	A30_X
X=0	104.47	-	-	-	-	-	-
X=5	-	100.78	99.16	97.87	91.18	86.40	79.49
X=10	-	99.37	99.52	96.29	94.41	88.84	80.19
X=15	-	99.34	99.52	96.42	94.25	89.55	81.10
X=20	-	101.19	99.33	98.11	92.58	87.56	83.29
X=25	-	101.03	101.09	96.44	93.37	88.10	86.27
X=30	-	99.02	98.85	97.55	93.13	90.59	91.18

Table 4. Stroke efficiency

St_e [-]	Column						
	SM	A05_X	A10_X	A15_X	A20_X	A25_X	A30_X
X=0	0.316	-	-	-	-	-	-
X=5	-	0.268	0.239	0.265	0.300	0.298	0.280
X=10	-	0.260	0.242	0.281	0.287	0.265	0.258
X=15	-	0.250	0.296	0.298	0.285	0.286	0.259
X=20	-	0.267	0.243	0.305	0.286	0.265	0.260
X=25	-	0.264	0.246	0.258	0.280	0.291	0.262
X=30	-	0.273	0.264	0.269	0.265	0.284	0.270

Table 5. Crash load efficiency

CLE [%]	Column						
	SM	A05_X	A10_X	A15_X	A20_X	A25_X	A30_X
X=0	31.37	-	-	-	-	-	-
X=5	-	30.39	29.72	31.03	35.12	36.91	39.19
X=10	-	30.49	29.71	32.39	33.30	34.29	37.67
X=15	-	30.07	31.99	33.10	33.26	35.04	37.32
X=20	-	30.20	29.80	32.84	33.84	34.79	36.38
X=25	-	30.12	29.42	31.34	33.32	35.87	35.19
X=30	-	31.12	30.81	31.41	32.73	34.51	33.63

i.e., at the base (bottom) and at the top end (where the crushing force was applied) of the column were examined.

The results demonstrate that the main factor influencing the crushing mode and, subsequently, energy absorption capability of the column is the depth of the dent. Three crushing modes were determined. The most desirable one is that initiated (triggered) by a local plastic mechanism situated exactly at the dent. This case produced the largest values of relative dent depth (A20-A30), regardless of the position of the dent. For this crushing mode the highest values of crash load efficiency (CLE) and the smallest values of peak crushing force (PCF- Table 3) were obtained, as illustrated in Fig. 16 and Fig.19. For the deepest dents, the relative increase in the CLE amounted to about 25% for the non-flawed column. The corresponding relative decrease in the PCF amounted to about 24%. For the deepest dents, where crushing mode III could be observed, the CLE indicator decreased with elevating the dent's position. At a smaller depth of the dent, for crushing modes I and II, the dent position has little effect on the energy absorption capability of the column.

The results of the parametric study indicate that the proposed design solution of an energy absorbing thin-walled structure is efficient. Thus, further research will be conducted on multi-objective crashworthiness optimization, taking into account a wide range of geometric and material parameters. In addition to this, an experimental validation of the numerical FE model and numerical results will be performed on columns with dents using a drop hammer rig.

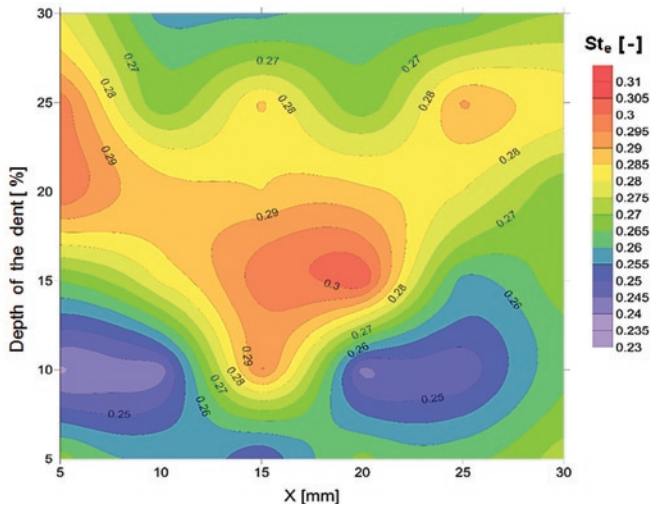


Fig. 18. Map of stroke efficiency in relative depth- position X plane

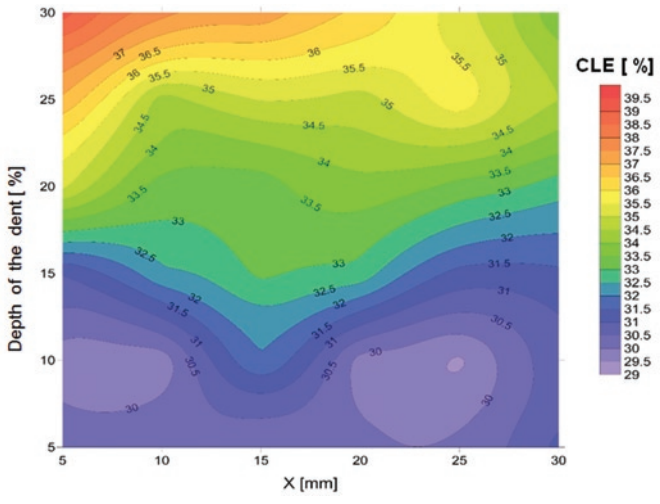


Fig. 19. Map of crash load efficiency in relative depth- position X plane

References

1. Abbasi M. et al. Multiobjective crashworthiness optimization of multi-cornered thinwalled sheet metal members, Thin Walled Struct 2015; 89: 31–41, https://doi.org/10.1016/j.tws.2014.12.009.
2. Abramowicz W. Thin-walled structures as impact energy absorbers. Thin Wall Struct. 2003; 41: 91 -109, https://doi.org/10.1016/S0263-8231(02)00082-4.
3. Alexander JM. An approximate analysis of the collapse of thin cylindrical shells under axial loading. Q J Mech Appl Math 1960; 13(1):10-15, https://doi.org/10.1093/qjmam/13.1.10.
4. Alghamdi AAA. Collapsible impact energy absorbers: an overview. Thin Wall Struct 2001; 39: 189-213, https://doi.org/10.1016/S0263-8231(00)00048-3.
5. Ali M, Ohioma E, Kraft F, Alam K. Theoretical, numerical and experimental study of dynamic axial crushing of thin walled pentagon and cross-shape tubes. Thin Wall Struct 2015; 94: 253-272, https://doi.org/10.1016/j.tws.2015.04.007.
6. Baroutaji A. et al., On the crashworthiness performance of thin-walled energy absorbers: recent advances and future developments. Thin-Walled Struct. 2017; 118: 137-163, https://doi.org/10.1016/j.tws.2017.05.018.
7. Chen W, Wierzbicki T. Relative merits of single-cell, multi-cell and foam-filled thin walled structures in energy absorption. Thin Wall Struct 2001; 39: 287-306, https://doi.org/10.1016/S0263-8231(01)00006-4.
8. Darvizeh A. et al. Low velocity impact of empty and foam filled circumferentially grooved thick-walled circular tubes. Thin Walled Struct 2017; 110: 97-105, https://doi.org/10.1016/j.tws.2016.09.002.
9. Estrada Q, Szwedowicz D, Majewski T, Martinez E, Rodriguez-Mendez A. Effect of quadrilateral discontinuity size on the energy absorption of structural steel profiles. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18 (2): 186–193, https://doi.org/10.17531/ein.2016.2.5.
10. Ferdynus M. An energy absorber in the form of a thin-walled column with square crosssection and dimples. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2013; 15:253-258.
11. Hanssen AG, Langseth M, Hopperstad OS. Static crushing of square aluminium extrusions with aluminium foam filler. Int J Mech Sci 1999; 41: 967-993, https://doi.org/10.1016/S0020-7403(98)00064-2.
12. Hanssen AG, Langseth M, Hopperstad OS. Static and dynamic crushing of square aluminium extrusions with aluminium foam filler. . Int J Mech Sci 2000; 24: 347-383, https://doi.org/10.1016/S0734-743X(99)00169-4.
13. Jones N. Structural Impact. Cambridge University Press 2003.
14. Jones N., Energy absorbing effectiveness factor, Int. J. of Impact Engineering 2010; 37: 754-765, https://doi.org/10.1016/j.ijimpeng.2009.01.008.
15. Kai Yang et al., Design of dimpled tubular structures for energy absorption. Thin-Walled Struct. , 2017, Thin-Walled Struct. 2017; 111: 1-8, https://doi.org/10.1016/j.tws.2016.11.004.
16. Lancaster ER, Palmer SC. Model Testing of Mechanically Damaged Pipes Containing Dents and Gouges. ASME Pressure Vessels & Piping Conference New York 1992; 235: 143–148.
17. Liu W. et al. Crushing behaviour and multi-objective optimization on the crashworthiness of sandwich structure with star-shaped tube in the center. Thin Wall Struct 2016; 108: 205-214, https://doi.org/10.1016/j.tws.2016.08.021.
18. Reddy S, Abbasi M, Fard M, Multi-cornered thin-walled sheet metal members for enhanced crashworthiness and occupant protection. Thin Walled Struct 2015; 94: 56-66, https://doi.org/10.1016/j.tws.2015.03.029.
19. Sharifi S. et al. Experimental investigation of bitubal circular energy absorbers under quasi-static axial load. Thin Wall Struct 2015; 89: 42-53, https://doi.org/10.1016/j.tws.2014.12.008.
20. Wang Q, Fan Z, Gui L. A theoretical analysis for the dynamic axial crushing behaviour of aluminium foam- filled hat sections. Int J Solids Struct 2006; 43: 2064-2075, https://doi.org/10.1016/j.ijsolstr.2005.06.011.
21. Yin H. et al. Crashworthiness optimization design for foam-filled multi-cell thin-walled structures. Thin Wall Struct 2015; 75: 8-17, https://doi.org/10.1016/j.tws.2013.10.022.
22. Zarei HR, Kroger M. Optimization of the foam-filled aluminium tube for crush box application. Thin Wall Struct 2008; 46: 214–221, https://doi.org/10.1016/j.tws.2007.07.016.
23. Zhe Yang et al. , Crushing behaviour of a thin-walled circular tube with internal gradient grooves fabricated by SLM 3D printing. Thin-Walled Struct. 2017; 111: 1-8, https://doi.org/10.1016/j.tws.2016.11.004.

Mirosław FERDYNUS
Lublin University of Technology,
Department of Machine Construction & Mechatronics,
ul. Nadbystrzycka 36, 20-618 Lublin, Poland

Maria KOTEŁKO
Lodz University of Technology,
Department of Strength of Materials
ul. Stefanowskiego 1/15 (A22), Łódź, Poland

Jan KRAL-jr
Technical University of Kosice
Faculty of Mechanical Engineering
Letna 9, 042 00 Kosice, Slovak Republic

Emails: m.ferdynus@pollub.pl, maria.kotelko@p.lodz.pl, kral.jan@tuke.sk

Article citation info:
ZIEJA M, WAŻNY M, STĘPIEŃ S. Outline of a method for estimating the durability of components or device assemblies while maintaining the required reliability level. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 260–266, http://dx.doi.org/10.17531/ein.2018.2.11.

Mariusz ZIEJA
Mariusz WAŻNY
Sławomir STĘPIEŃ

OUTLINE OF A METHOD FOR ESTIMATING THE DURABILITY OF COMPONENTS OR DEVICE ASSEMBLIES WHILE MAINTAINING THE REQUIRED RELIABILITY LEVEL

ZARYS METODY SZACOWANIA TRWAŁOŚCI ELEMENTÓW LUB ZESPOŁÓW URZĄDZEŃ Z ZACHOWANIEM WYMAGANEGO POZIOMU NIEZAWODNOŚCI*

The paper includes a probabilistic method for evaluating the durability of components and device assemblies which operate under the impact of destructive processes. As a result of these processes, wear that causes deterioration of their cooperation conditions occurs. It is assumed that a component operates reliably when the wear does not exceed the acceptable (limit) values. In mathematical terms, this method is based on a differential equation, after the transformation of which, it is possible to obtain the Fokker-Planck type partial differential equation. The specific solution of this equation allows for obtaining the density function of the probability wear in the normal distribution form. The paper presents two methods for determining the durability. The first one involves the application of the wear density function, and the second one consists in determining the probability density function of the time of reaching the acceptable state, and its use in order to determine the component or assembly durability. The paper presents a numerical example on the aircraft technology operation process.

Keywords: reliability, durability, density function acceptable state, ageing, wear.

Praca zawiera probabilistyczną metodę oceny trwałości elementów lub zespołów urządzeń pracujących w warunkach oddziaływania procesów destrukcyjnych. W wyniku działania tychże procesów następuje zużywanie powodujące pogorszenie warunków ich współpracy. Przyjmuje się, że element pracuje niezawodnie, gdy zużycie nie przekracza wartości dopuszczalnych (granicznych). Metoda od strony matematycznej bazuje na równaniu różnicowym z którego po przekształceniu otrzymuje się równanie różniczkowe cząstkowe typu Fokkera-Plancka. Z rozwiązania szczególnego tego równania otrzymuje się funkcję gęstości prawdopodobieństwa zużywania w postaci rozkładu normalnego. W pracy przedstawione są dwa sposoby wyznaczania trwałości. Pierwszy polega na wykorzystaniu funkcji gęstości zużywania a drugi na wyznaczeniu funkcji gęstości prawdopodobieństwa czasu osiągania stanu dopuszczalnego i zastosowanie jej do wyznaczenia trwałości elementu lub zespołu. W pracy przedstawiono przykład liczbowy dotyczący procesu eksploatacji techniki lotniczej.

Słowa kluczowe: niezawodność, trwałość, funkcja gęstości stan dopuszczalny, starzenie, zużywanie

1. Introduction

In the available literature, it is possible to find a number of papers, which demonstrate the problem of the impact of the external environment, ageing and wear processes on the technical system functioning [4, 9, 13, 16, 17, 21]. Due to technical advancement and a high degree of integration of the devices used on the board of military aircraft, the development of optimal operation models is a complex task. The methods for evaluating the reliability and durability of aviation equipment based on a change in diagnostic parameters are extremely useful within this area [6, 7, 8, 12, 15, 20].

This paper includes a probabilistic method for evaluating the durability of components and the assemblies of the device that operates under the impact of ageing processes (corrosive, wear and other) in the aircraft devices [15, 18, 19]. The technical condition of some aviation equipment can be assessed with the use of diagnostic parameters. This assessment requires knowledge of limit (acceptable) values, for which it is considered that the device or assembly is in the state of usability.

In the offered durability assessment model, the following assumptions are adopted:

– the device's technical condition is defined by one diagnostic parameter "z" in the form of the parameter deviation from the nominal value:

$$z = \left| X - X^{norm} \right|, \tag{1}$$

where:

X – current value of the diagnostic parameter,
X^{norm} – nominal value of the diagnostic parameter;

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

– change in the deviation value of the diagnostic parameter takes place in the entire operation period (operation and standstill);
– "z" parameter is non-decreasing, because it is determined by the absolute value of the difference of the present and nominal values;
– increase speed of the diagnostic parameter deviation in case of random changes can be described by the following relationship:

$$\frac{dz}{dt} = c\,, \tag{2}$$

where:

c – random variable which characterises the component's susceptibility to ageing changes depending on its features and operating conditions,
t – calendar time.

2. Method for estimating the durability of the device component with the use of the density function of the diagnostic parameter deviation

2.1. Determination of the deviation density function taking into account the relationship (1)

The dynamics of changes in "z" deviation value in random terms will be characterised by the following differential equation:

$$U_{z,t+\Delta t} = (1-P)U_{z,t} + PU_{z-\Delta z,t}\,, \tag{3}$$

where:

$U_{z,t}$ – probability of the fact that in the moment of t, the diagnostic parameter value adopts z value;
P – probability of the event that the random wear occurs and that in the time interval of Δt, the deviation value will be increased by Δz value;
Δz – deviation increase.

In case, when $P=1$ equation (3) in the function notation will adopt the following form:

$$u(z,t+\Delta t) = u(z-\Delta z,t)\,. \tag{4}$$

The equation (4) has the following form: probability of the fact that in the moment of $t+\Delta t$, the deviation value will be z is equal to the probability of the fact that in the t moment, the deviation value was equal to z-Δz. It means that along with the probability equal to unity, in the time interval of Δt, the deviation will be increased by Δz value.

The equation (4) is transformed into the partial differential equation. Therefore, the following approximations are adopted [1,2]:

$$u(z,t+\Delta t) = u(z,t) + \frac{\partial u(z,t)}{\partial t}\Delta t\,, \tag{5}$$

$$u(z-\Delta z,t) = u(z,t) - \frac{\partial u(z,t)}{\partial z}\Delta z + \frac{1}{2}\frac{\partial^2 u(z,t)}{\partial z^2}(\Delta z)^2\,. \tag{6}$$

By using (5) and (6), the equation (4) adopts the following form:

$$\frac{\partial u(z,t)}{\partial t} = -b\frac{\partial u(z,t)}{\partial z} + \frac{1}{2}a\frac{\partial^2 u(z,t)}{\partial z^2}\,, \tag{7}$$

where:

$b=E[c]$ – average increase in the diagnostic parameter's deviation value per time unit;
$a=E[c^2]$ – average increase square of the diagnostic parameter's deviation per time unit.

We are searching for the solution of a particular equation (7), the one, which at $t\to 0$ is coergent to the so-called Dirac function, i.e. $u(z,t)\to 0$ for $z\neq 0$ and $u(0,t)\to +\infty$, but in a way that the integral of u function is equal to "1" for all $t>0$.

The equation solution (7) adopts the following form for the above specified condition [3, 11, 14]:

$$u(z,t) = \frac{1}{\sqrt{2\pi A(t)}}e^{-\frac{(z-B(t))^2}{2A(t)}}\,, \tag{8}$$

where:

$$B(t) = \int_0^t b\,dt = bt = \bar{c}t, \qquad A(t) = \int_0^t a\,dt = at = \bar{c}^2 t\,.$$

The value of 0 in lower limits of the integrals means the adopted initial moment of time, according to which the dynamics of changes in the diagnostic parameter's value is considered – it can be, e.g. the moment of putting a given device into operation.

The density function (8) of the diagnostic parameter's deviation increase can be used for assessing the reliability of the considered device component.

2.2. Determination of reliability and durability of the component or device assembly

By having a specific density function, it is possible to record the relationship on reliability and durability due to the time of the parameter's deviation increase to the limit value. The formula adopts the following form:

$$R(t) = \int_{-\infty}^{z_d} u(z,t)\,dz, \tag{9}$$

where:

$u(z,t)$ – density function specified by the relationship (8);
z_d – acceptable value of the diagnostic parameter's deviation due to safety;
t – calendar time of the device operation.

Figure 1 presents a diagram of the density function course and a way of determining the reliability and durability.

The relationship (9) taking into account (8), adopts the following form:

$$R(t) = \int_{-\infty}^{z_d} \frac{1}{\sqrt{2\pi at}}e^{-\frac{(z-bt)^2}{2at}}\,dz\,. \tag{10}$$

By assuming the minimum, required value of R^* reliability, it is possible to determine t^*time, after which the reliability will decrease below the required level. The time t^* can be treated as the durability of a given component for the required, acceptable reliability value.

In this case, it is possible to obtain:

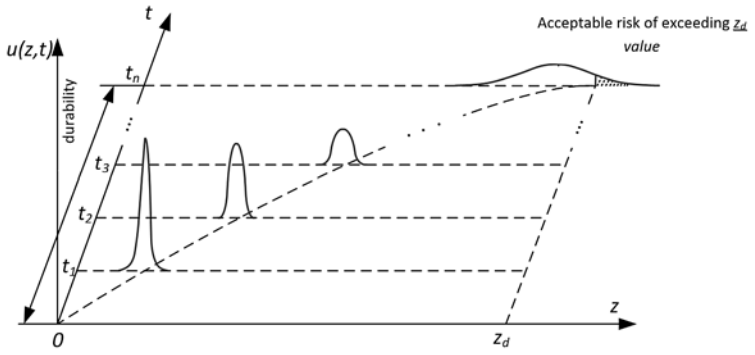


Fig. 1. Diagram of changes in the density function form

$$R^* = \int_{-\infty}^{z_d} \frac{1}{\sqrt{2\pi at^*}} e^{-\frac{\left(z-bt^*\right)^2}{2at^*}} dz \,. \tag{11}$$

3. Method for estimating the durability with the use of the density function of the time exceeding the acceptable (limit) state

3.1. Determination of the time distribution of exceeding the acceptable (limit) state

The probability of exceeding the acceptable (limit) value by the diagnostic parameter with the use of the density function of changes in the diagnostic parameter's deviation (8) can be presented in the following form:

$$Q(t;z_d) = \int_{z_d}^{\infty} \frac{1}{\sqrt{2\pi at}} e^{-\frac{(z-bt)^2}{2at}} dz \,. \tag{12}$$

The density function of the time distribution of the first transition beyond the acceptable value z_d adopts the following form:

$$f(t) = \frac{\partial}{\partial t} Q(t;z_d) = \frac{\partial}{\partial t} \int_{z_d}^{\infty} \frac{1}{\sqrt{2\pi at}} e^{-\frac{(z-bt)^2}{2at}} dz \,. \tag{13}$$

Thus,

$$f(t) = \int_{z_d}^{\infty} \left\{ \frac{\partial}{\partial t} \left[\frac{1}{\sqrt{2\pi at}} e^{-\frac{(z-bt)^2}{2at}} \right] \right\} dz \,. \tag{14}$$

By assuming (8) definition, it is possible to obtain:

$$f(t) = \int_{z_d}^{\infty} \left\{ \frac{\partial}{\partial t} u(z,t) \right\} dz \,. \tag{15}$$

Furthermore, a derivative after the function time (8), adopts the following form:

$$\frac{\partial}{\partial t}\left[u(z,t)\right] = u(z,t)\left(\frac{z^2 - b^2t^2 - at}{2at^2}\right). \tag{16}$$

The relationship (16) was substituted to (14):

$$f(t) = \int_{z_d}^{\infty} \left[u(z,t)\left(\frac{z^2 - b^2t^2 - at}{2at^2}\right)\right] dz \,. \tag{17}$$

The primary function for the integrand in the relationship (17) is searched for. It is expected that the function in the form of:

$$w(z,t) = u(z,t)\theta(z,t),$$

is a primary function for the integrand of the relationship (17), where $\theta(z,t)$ is a sought unknown function.

That is:

$$\frac{\partial}{\partial z}\left[u(z,t)\theta(z,t)\right] = u(z,t)\left(\frac{z^2 - b^2t^2 - at}{2at^2}\right).$$

After transformations, the following equation is obtained:

$$\frac{\partial\theta(z,t)}{\partial z} - \frac{z-bt}{at}\theta(z,t) = \frac{z^2 - b^2t^2 - at}{2at^2} \,. \tag{18}$$

Homogeneous equation:

$$\frac{\partial\theta(z,t)}{\partial z} - \frac{z-bt}{at}\theta(z,t) = 0 \,.$$

Solution of the homogeneous equation:

$$\theta_0(z,t) = Ce^{\frac{z^2-2btz}{2at}},$$

where: C – arbitrary constant

The expected specific solution of the homogeneous equation has the following form:

$$\theta_s(z,t) = -\frac{z+bt}{2t} \,.$$

It was checked that the equation (18) fulfils the above solution. The general solution of the homogeneous equation:

$$\theta(z,t) = Ce^{\frac{z^2-2btz}{2at}} - \frac{z+bt}{2t} \,.$$

That is the sought primary function of the integral (17) has the following form:

$$w(z,t)=u(z,t)\left[Ce^{\frac{z^2-2btz}{2at}}-\frac{z+bt}{2t}\right].$$

Thus, by calculating the integral (17) in the specified limits, it is possible to obtain:

$$f(t)=u(z,t)\left[Ce^{\frac{z^2-2btz}{2at}}-\frac{z+bt}{2t}\right]_{z_d}^{\infty}=Cu(z,t)e^{\frac{z^2-2btz}{2at}}\Big|_{z_d}^{\infty}-u(z,t)\frac{z+bt}{2t}\Big|_{z_d}^{\infty}=$$

$$=C\frac{1}{\sqrt{2\pi at}}e^{-\frac{b^2t}{2a}}\Big|_{z_d}^{\infty}-u(z,t)\frac{z+bt}{2t}\Big|_{z_d}^{\infty}=0-0+u(z_d,t)\frac{z_d+bt}{2t}$$

$$f(t)=\frac{z_d+bt}{2t}u(z_d,t). \tag{19}$$

The relationship (19) determines the density function of the time of the first transition of the acceptable (limit) state by the diagnostic parameter's deviation. It should be checked, whether the function (19) is a density function of time of reaching the acceptable (limit) state. The function has the following form:

$$f(t)=\frac{z_d+bt}{2t}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z_d-bt)^2}{2at}}. \tag{20}$$

The function (20) should meet the condition:

$$\int_0^{\infty}f(t)dt=1. \tag{21}$$

In order to demonstrate the validity (21), the following justification is presented:

$$\int_0^{\infty}\frac{z_d+bt}{2t}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z_d-bt)^2}{2at}}dt=1. \tag{22}$$

In order to calculate the integral that occurs in the formula (22), the following substitution is used:

$$w=\frac{z_d-bt}{\sqrt{at}}\Rightarrow dt=-\frac{2t\sqrt{at}}{z_d+bt}dw. \tag{23}$$

Transformation of the limits of integration:

$$t=0\Rightarrow w=\infty,$$
$$t=\infty\Rightarrow w=\lim_{t\to\infty}\frac{z_d-bt}{\sqrt{at}}=\lim_{t\to\infty}\frac{-2b\sqrt{at}}{a}=-\infty. \tag{24}$$

After substituting to the output integral, it is possible to obtain:

$$\int_0^{\infty}\frac{z_d+bt}{2t}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z_d-bt)^2}{2at}}dt=-\int_{\infty}^{-\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw \tag{25}$$

The above integral is an integral of $N(0,1)$ normal distribution in the limits from -∞ to +∞ and is equal to unity. On this basis, it can be concluded that:

$$\int_0^{\infty}\frac{z_d+bt}{2t}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z_d-bt)^2}{2at}}dt=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw. \tag{26}$$

3.2. Evaluation of the durability of selected components of the aircraft construction with the use of the time distribution of obtaining the acceptable state

The formula for the aircraft's structural component reliability adopts the following form:

$$R(t)=1-\int_0^t f(\tau)d\tau, \tag{27}$$

where the density function $f(t)$ is determined by the following formula (19).

However, the unreliability of the aircraft's structural component can be determined on the basis of the following relationship:

$$Q(t)=\int_0^t\frac{z_d+b\tau}{2\tau}\frac{1}{\sqrt{2\pi a\tau}}e^{-\frac{(z_d-b\tau)^2}{2a\tau}}d\tau. \tag{28}$$

The integral occurring in the relationships (27) and (28) must be transformed to the more convenient form:

$$\int_0^t\frac{z_d+b\tau}{2\tau}\frac{1}{\sqrt{2\pi a\tau}}e^{-\frac{(z_d-b\tau)^2}{2a\tau}}d\tau=\left|\begin{array}{ll}w=\frac{z_d-b\tau}{\sqrt{a\tau}} & \tau=0\Rightarrow w=\infty\\ d\tau=-\frac{2\tau\sqrt{a\tau}}{z_d+b\tau} & \tau=t\Rightarrow w=\frac{z_d-bt}{\sqrt{at}}\end{array}\right|=-\int_{\infty}^{\frac{z_d-bt}{\sqrt{at}}}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw.$$

After changing the limits of integration, it is possible to obtain:

$$\int_0^t\frac{z_d+b\tau}{2\tau}\frac{1}{\sqrt{2\pi a\tau}}e^{-\frac{(z_d-b\tau)^2}{2a\tau}}d\tau=\int_{\frac{z_d-bt}{\sqrt{at}}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw. \tag{29}$$

The reliability of a given component will adopt the following form:

$$R(t)=1-\int_{\frac{z_d-bt}{\sqrt{at}}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw, \tag{30}$$

or

$$R(t)=\int_{-\infty}^{\frac{z_d-bt}{\sqrt{at}}}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw. \tag{31}$$

The integral occurring in the formula (31) is a value of $N(0,1)$ normal distribution function for the argument occurring in the upper limit of integration. Again, by assuming the required minimum value of R^* reliability, it is possible to determine t^* durability.

$$R^* = \int_{-\infty}^{\frac{z_d - bt^*}{\sqrt{at^*}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw. \quad (32)$$

The use of (11) or (32) formula in the calculation requires estimation of the values of *a* and *b* coefficients. This estimation is carried out on the basis of the data obtained from the aircraft operation process.

4. Numerical example

In order to determine the durability of the considered component, it is important to determine (estimate) the values of *a* and *b* constants. Therefore, it is assumed that the observation of the tested device in the operation process results in the provision of data on the increase of the diagnostic parameter's deviation value in the form of:

$$\left[(z_0,t_0),(z_1,t_1),(z_2,t_2),\ldots,(z_n,t_n)\right]. \quad (33)$$

The best method for determining "*b*" and "*a*" values for the held data is a method that uses a likelihood function. Its form in the general case can be presented as the relationship:

$$L = \prod_{k=0}^{n-1} g(t_k, z_k, \theta_1, \theta_2, \ldots, \theta_m), \quad (34)$$

where:

$g(t_k, z_k, \theta_1, \theta_2, \ldots, \theta_m)$ – density function of the total probability of *z* variable;

$(\theta_1, \theta_2, \ldots, \theta_m)$. – density function parameters;

z_k – measured wear values of *z* parameter respectively in the moments of time $(t_1, t_2, \ldots, t_k)$.

Finding $(\theta_1^*, \theta_2^*, \ldots, \theta_m^*)$ estimates of unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$ with the use of a maximum likelihood method consists in solving the equations in the form of:

$$\frac{\partial lnL}{\partial \theta_j} = 0, \quad (35)$$

where:

j=1,2,…,m;

m – number of parameters characterising the wear process of a given technical object.

In this case, b^* and a^* estimates of unknown *b* and *a* parameters with the use of the maximum likelihood method consists in solving the system of equations:

$$\begin{cases} \frac{\partial lnL}{\partial b} = 0 \\ \frac{\partial lnL}{\partial a} = 0 \end{cases}. \quad (29)$$

By solving the system of equations (29), b^* and a^* are found.

$$b^* = \frac{z_n}{t_n}, \quad (36)$$

$$a^* = \frac{1}{n} \sum_{k=0}^{n-1} \frac{\left[(z_{k+1} - z_k) - b^*(t_{k+1} - t_k)\right]^2}{(t_{k+1} - t_k)}. \quad (37)$$

The component, which was chosen for a numerical example is 12-SAM-28 aircraft battery. Figure 2 shows a change in the time of the averaged battery capacity for held data.

In accordance with the relationship (1), the absolute value of the capacity difference and its nominal value were adopted as "*z*" diagnostic parameter. The change in time of "*z*" parameter was presented in Figure 3.

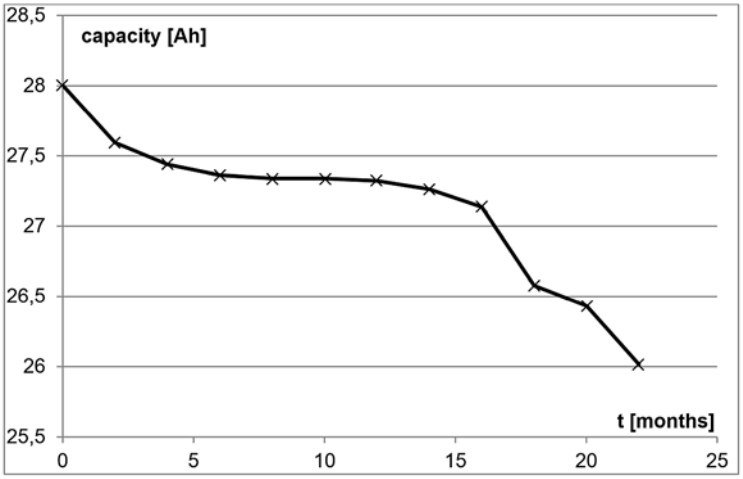


Fig. 2. The course of changes in the averaged capacity of 12-SAM-28 battery

Thus, holding the data describing the values of the diagnostic parameter in the form of $\left[(z_0,t_0),(z_1,t_1),(z_2,t_2),\ldots,(z_n,t_n)\right]$, based on (36) and (37) formulas, the values of the density function coefficients were determined:

$$b^*=0{,}09, \quad a^*=0{,}015. \quad (38)$$

The parameter z_d was determined with the use of technical documentation used for the implementation of maintenance works, in which the information on the acceptable value of the capacity of batteries was provided.

Therefore, by holding the values of parameters b_ε^*, a_ε^*, z_d, they were substituted to (11) or (32) equations by determining the relationship of t^* time on R^* probability – Figure 4. In both cases (relationship (11) or (32)) the same course was obtained.

By assuming the minimum value of R^*=0.99 reliability, the time, to which the diagnostic parameter deviation will not exceed the limit state, in accordance with the assumed probability, was determined:

$$T=63 \text{ [months]}. \quad (39)$$

The obtained value (39) can be used in the technical maintenance depending on the adopted strategy of maintenance. On the basis of the above methodology, it is possible to determine further periods, in which the control of the device diagnostic parameter should be carried out [5, 10].

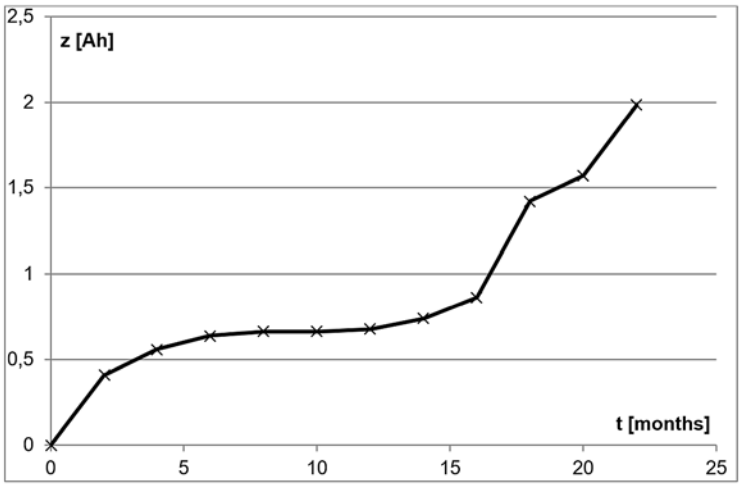


Fig. 3. Change in time of "z" parameter for 12-SAM-28 battery

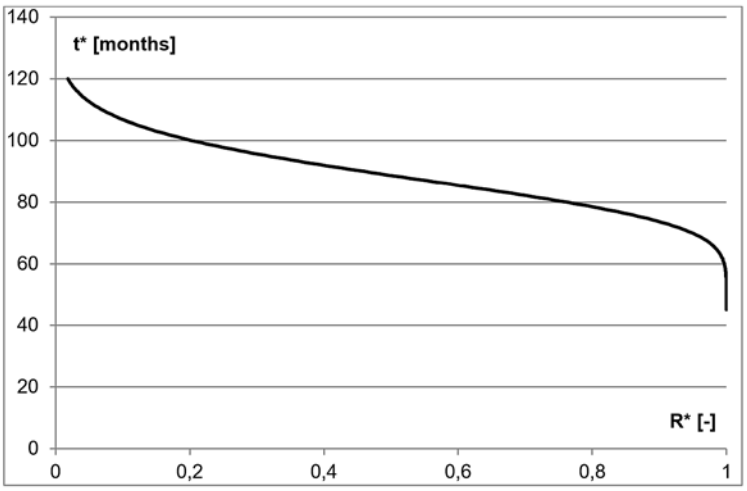


Fig. 4. Relationship of projected t^* durability on R^* reliability

5. Final remarks

In this paper, an overview of the method for estimating the durability of components or assemblies, when the increase speed of changes was of random nature, was presented. However, the method of this change was described by the following simple relationship:

$$\frac{dz}{dt} = c \,,$$

where c was a random variable determining the possibility of the parameter's deviation increase.

It is possible to generalise this method, when the speed of the deviation increase will be described by the following relationships:

$$\frac{dz}{dt} = cz \,, \tag{40}$$

$$\frac{dz}{dt} = ct^{\alpha-1} \,. \tag{41}$$

In the first case, the increase speed of changes will be of random nature similar to the exponential one. In the second case, the increase nature of changes will be similar to the intensity of damage in the Weibull distribution.

In summary, it can be concluded that the presented method seems to be correct and right, and allows to analyse the device technical condition due to the nature of changes in the values of diagnostic parameters. The presented calculation example allowed to carry out the verification of the developed model, and emphasised the developed method's application advantages. This method may be useful in further works on the improvement of both the operational process and the method of using the aircraft with the use of its on-board systems, allowing for determining the time of the device's staying in the state of usability.

Furthermore, the presented method, owing to its universal nature, can be successfully used in order to specify the residual life of any technical object, the technical condition of which is determined on the basis of the analysis of the diagnostic parameters' values.

In this paper, the presented method can be further improved and extended to other cases of increase in random changes of the exponential type. It seems that it can be used for assessing the reliability of mechanical components, in case of considering the propagation of fatigue cracks in the components subjected to the random load, and in case of using the Paris formula in order to specify the crack velocity.

References

1. DeLurgio SA. Forecasting principles and applications. University of Missouri-Kansas City: Irwin/McGraw-Hill, 1998.
2. Franck TD. Nonlinear Fokker-Planck Equations. Fundamentals and Applications. Berlin Heildenberg: Springer-Verlag, 2005.
3. Grasman J, Herwaarden OA. Asymptotic Methods for the Fokker-Planck Equation and the Exit Problem in Applications. Berlin Heildenberg: Springer-Verlag, 1999, https://doi.org/10.1007/978-3-662-03857-4.
4. Idziaszek Z, Grzesik N. Object characteristics deterioration effect on task realizability – outline method of estimation and prognosis. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (3): 433–440.
5. Kinnison H, Siddiqui T. Aviation Maintenance Management. The McGraw-Hill Companies, Inc. 2013.
6. Knopik L, Migawa K. Multi-state model of maintenance policy. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (1): 125–130, https://doi.org/10.17531/ein.2018.1.16.
7. Knopik L, Migawa K. Wdzięczny A. Profit optimization in maintenance system, Polish Martime Research, 2016, 1(89): 193-98.
8. Kołowrocki K, Soszyńska Budny J. Reliability and Safety of Complex Technical Systems and Processes. Springer 2011, https://doi.org/10.1007/978-0-85729-694-8.
9. McPherson JW. Reliability physics and engineering. New York: Springer, 2010, https://doi.org/10.1007/978-1-4419-6348-2.
10. Narayan V. Effective Maintenance Management. New York: Industrial Press Inc., 2012.
11. Pham H. Handbook of Engineering Statistics. London: Springer-Verlag 2006, https://doi.org/10.1007/978-1-84628-288-1.
12. Rasuo B., Duknic G. Optimization of the aircraft general overhaul process. Aircraft engineering and aerospace technology 2013; 85 (5):

343-354, https://doi.org/10.1108/AEAT-02-2012-0017.
13. Restel F. The Markov reliability and safety model of the railway transportation system. Safety and reliability: methodology and applications: proceedings of the European Safety and Reliability Conference, ESREL 2014, 14-18 September, 2015, Wrocław, Poland. CRC Press/ Balkema: 303-311.
14. Risken H. The Fokker-Planck Equation. Methods of Solution and Applications. Berlin Heildenberg: Springer Verlag, 1984, https://doi.org/10.1007/978-3-642-96807-5.
15. Tan CM, Singh P. Time evolution degradation physics in high power white LEDs under high temperature-humidity conditions. IEEE Transactions on Device and Materials Reliability 2014; 14(2): 742-750, https://doi.org/10.1109/TDMR.2014.2318725.
16. Ułanowicz L. Modelling of a process, which causes adhesive seizing (tacking) in precise pairs of hydraulic control devices. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18 (4): 492-500, https://doi.org/10.17531/ein.2016.4.3.
17. Valis D, Koucky M, Zak L. On approaches for non-direct determination of system deterioration. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 1:33-41.
18. Wang P, Tang Y, Baeb SJ, He Y. Bayesian analysis of two-phase degradation data based on change-point Wiener process. Reliability Engineering & System Safety 2018; 170: 244-256, https://doi.org/10.1016/j.ress.2017.09.027.
19. Wang YS, Zhang CH, Zhang SF, Chen X, Tan YY. Optimal design of constant stress accelerated degradation test plan with multiple stresses and multiple degradation measures. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2015; 229(1): 83-93, https://doi.org/10.1177/1748006X14552312.
20. Woch M. Reliability analysis of the PZL-130 Orlik TC-II aircraft structural component under real operating conditions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (2): 287–295, https://doi.org/10.17531/ein.2017.2.17.
21. Zurek J, Tomaszek H. Zieja M. Analysis of structural component's lifetime distribution considered from the aspect of the wearing with the characteristic function applied. Safety, reliability and risk analysis: Beyond the horizon. Amsterdam: Balkema 2014, 2597-2602.

Mariusz ZIEJA
Air Force Institute of Technology
ul. Księcia Bolesława 6, 01-494 Warsaw 96, Poland

Mariusz WAŻNY
Sławomir STĘPIEŃ
Faculty of Mechatronics and Aerospace
Military University of Technology
ul. Kaliskiego 2, 00-908 Warsaw 49, Poland

E-mails: mariusz.zieja@itwl.pl, mwazny@wat.edu.pl, sstepien@wat.edu.pl

Article citation info:
ZARNIK MS, NOVAK F, PAPA G. Sensors in proactive maintenance – a case of LTCC pressure sensors. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 267–272, http://dx.doi.org/10.17531/ein.2018.2.12.

Marina SANTO ZARNIK
Franc NOVAK
Gregor PAPA

SENSORS IN PROACTIVE MAINTENANCE – A CASE OF LTCC PRESSURE SENSORS

CZUJNIKI STOSOWANE W KONSERWACJI PROAKTYWNEJ – PRZYPADEK CERAMICZNYCH CZUJNIKÓW CIŚNIENIA WYKONANYCH W TECHNOLOGII LTCC

Sensors are a vital component part of any process-controlled system. Even though designed to properly operate at required conditions within the whole lifetime, all sensors exhibit some level of drift with time. When selecting the sensors for implementation in a system proactive maintenance their ageing in specific operating conditions should be considered as an important issue. Here we focus on thick-film piezoresistive sensors based on low temperature cofired ceramic (LTCC) and discuss their ageing in different regimes of operations. Frequent overloading and particularly with limit overpressures can result in observable drifts and unacceptable scattering from the calibrated characteristics. For the sensors operating in the water the overloads are even more critical. Moreover, under the regime with frequent overloads, some non-critical, intrinsic defects in the sensing structure, which normally do not affect the characteristics and are non-detectable by the output tests in serial production may develop into critical defects that shorten the sensor lifetime.

Keywords: *sensor ageing, pressure cycles, overpressure, low-frequency noise, system maintenance.*

Czujniki stanowią istotny komponent każdego systemu kontrolowanego przez proces. Chociaż czujniki są zaprojektowane tak, aby prawidłowo działały w wymaganych warunkach w całym okresie eksploatacji, wszystkie wykazują jednak pewien poziom dryfu w czasie. Wybierając czujniki do wdrożenia w proaktywnej konserwacji systemu, należy koniecznie rozważyć ich starzenie się w określonych warunkach pracy. Przedmiotem artykułu są grubowarstwowe czujniki piezorezystancyjne wykonane z ceramiki technologią LTCC (niskotemperaturowej ceramiki współwypalanej) oraz ich starzenie się w różnych trybach działania. Częste przeciążanie, zwłaszcza przy nadciśnieniu granicznym, może powodować dryf i niedopuszczalny rozrzut wskazań w stosunku do charakterystyk wzorcowych. W przypadku czujników pracujących w wodzie, przeciążenia mają jeszcze bardziej krytyczny charakter. Co więcej, w trybie pracy z częstymi przeciążeniami, niektóre niekrytyczne wady wewnętrzne w strukturze sensorowej, które normalnie nie mają wpływu na charakterystykę czujnika i są niewykrywalne w badaniach kontrolnych wyrobu gotowego w produkcji seryjnej, mogą przeobrażać się w wady krytyczne, które skracają cykl życia czujnika.

Słowa kluczowe: *starzenie się czujników, cykle ciśnienia, nadciśnienie, szum o niskiej częstotliwości, eksploatacja systemu.*

1. Introduction

Pressure sensors are the third most widely used category of physical sensors in the system maintenance [1]. Beside system control the pressure measurements can be used to reveal physical changes in a system. If the pressure values go outside the expected range, there is a possibility of damaging system parts. In safety-critical systems special measures are taken to provide robust and reliable operation even in the case of sensor failures [7]. Such solutions either introduce hardware or software redundancy, or make the monitoring more complex by enabling it to work better with incomplete data. Not all sensors can be made redundant because of space and cost constraints and the impact of individual sensors on system operation is analysed in order to identify their significance [5]. For the systems that are not necessarily safety-critical but still require reliable operation various possible maintenance scenarios with respect to both reliability and economic criteria are considered. For example, a method used to study the effect of equipment aging under different maintenance strategies is described in [13].

Accurate and reliable sensor measurements are prerequisite for efficient system maintenance, which is also one of the goals of the ECSEL JU project MANTIS – Cyber Physical System based Proactive Collaborative Maintenance [11], and the work reported in this paper. Proactive maintenance (PM) is a maintenance strategy which attempts to anticipate machine failures and similar problems and provides solutions before they occur [3]. The proactive approach benefits from the traditional preventive and predictive maintenance practice and upgrades them by root cause analysis, and predictive algorithms based on cyber physical system models [2, 9, 10]. The proactive maintenance can yield considerable savings over conventional predictive/preventive maintenance programs [4, 14].

Pressure sensors enable monitoring of operating conditions and the wear-out of the critical parts of the mechanical system for which proactive maintenance is performed. Yet, it should be noted that the sensors themselves are not ideal components, but physical devices including sensing elements, electronic components, connections and housings that are aging and degrading [12, 15], and in general may need to be maintained. To improve the reliability of a system, it is essential to ensure reliable operation of the sensors involved, i.e., the correct interpretation of their responses in the different periods of their life-cycles. In this regard, it is prudent to track and monitor their health conditions. In the presented case study, we concentrate on how the operating mode affects their ageing and identify measures that could be taken in defining the maintenance strategy.

The harsh operating conditions accelerate the ageing of sensors. The designers should be aware of this when selecting the appropriate sensors for the target application. Application notes and specifications normally provide data related to standard testing procedures. However, the actual circumstances in practice may impose specific conditions that system designers should be aware of when selecting the appropriate sensors for the target application.

The aging of sensors is an important issue, yet the reports on this are relatively few and mainly for the sensors based on the mature and well-established sensor technologies. In this paper, we present our findings related to the aging of the thick-film piezoresistive pressure sensors realised in Low Temperature Cofired (LTCC) which is a promising technology for wide range of sensor applications [6, 8]. Although a narrow field, the presented results may be helpful to raise awareness in proactive maintenance system design.

2. Pressure sensors in PM

Regardless of the specific applications and the complexity of the pressure sensor realizations the sensor itself can be considered as micro/mezzo system which, depending on the physical principle, includes various electromechanical structures (diaphragms, consoles) and the electronics. Pressure sensors can also include sensors of other physical quantities, for example temperature sensors that are used for analog temperature compensation. The aging of a sensor system is highly dependent on the operating conditions (environment, intensity and frequency of loadings) so it would be a good strategy to use PM for the sensor systems themselves. If in addition to the pressure some other variables are measured, such as temperature and relative humidity, such additional information can be used together with the pressure readings for the maintenance of the pressure sensors. The pressure sensors are systems composed of moving parts and functional materials that deteriorate over time. Acquisition and analysis of the sensor readings enables early detection of soft faults resulting from this deterioration.

2.1. LTCC-based pressure sensors

Ceramic pressure sensors are in general used in harsh environments and as such can be appropriate for integration in the system proactive maintenance. In this work we consider the thick-film piezoresistive pressure sensors in full Wheatstone bridge configuration realized in LTCC technology. The basic ceramic structure with a thin sensing diaphragm over the pressure cavity and with the integrated channels is made of DuPont 951 green tape. The bridge resistors are made of Du Pont 2041 thick-film resistor material. One representative sample of such sensor and the schematic representation of its cross section are presented in Fig. 1. The basic sensor embodiment can be completed with the readout electronics and may be installed in a housing. Depending on the application the sensors can be protected against harmful effects of the environment and measured media by different protective coatings. In the following we focus on uncompensated sensors without the readout electronics and without the housing. Although housing and protective coating may change the sensor characteristics the trends presented in this paper play important role in practice.

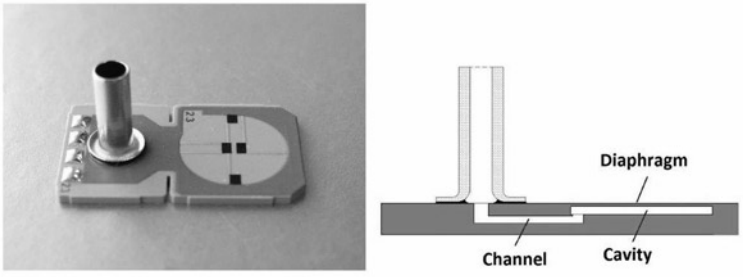


Fig. 1. Example LTCC pressure sensor and its cross section (not to scale)

The typical sensitivity (S) of the 100-mbar sensors considered in our case is 14 μV/V/mbar and the long-term stability is better than 0.2% per year. The stability of the offset voltage ($\Delta V = V_{off} - V_{off0}$) at the supply voltage of 5V is less than 100 μV per year. The typical response of the sensor measured in the pressure range from -100 mbar to 100 mbar at 25°C and the relative humidity (RH) of 40% is presented in Fig. 2.

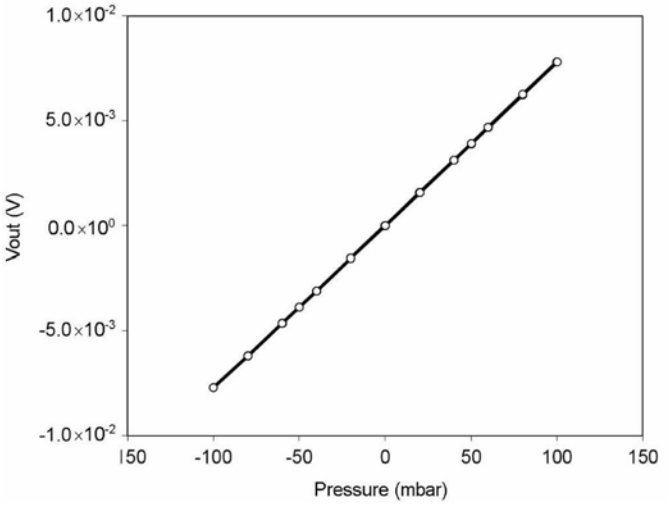


Fig. 2. Typical response of the 100-mbar sensor measured at 25°C and 40% RH

Previous studies [16, 17] have shown that the operating conditions in the regime of frequent overloading have non-negligible effects on the aging of such sensors. For efficient maintenance, accurate readings of such sensors including appropriate monitoring of the sensors performances and the operation conditions would be beneficial.

3. Ageing of LTCC pressure sensors

This section presents an additional analysis of the experimental results of the earlier case studies (i.e., the feasibility study and the early development phase of several LTCC-based pressure sensors) supplemented by the finding of further experimentation with the prototypes aimed at evaluation of their ageing performances.

It is normally to expect that sensor characteristics change over the time due to the aging of functional materials, even if they are not in contact with harsh media that accelerates the aging process. The aging is therefore highly dependent on the conditions of operation and storage, imposed by the environment (temperature, medium, continuous or randomly changing environment), location and regime of loading. In order to reduce the overall cost of proactive maintenance one might be tempted to employ cheap sensors with low cost housing. However, to assure accurate operation in a long lifetime they should be appropriately protected and maintained.

The LTCC sensors are subject to aging due to the aging of the functional materials and wear-out in long-term operation. The aging of the sensor is also affected by possible intrinsic defects in the material / structure, which in normal operation cannot be identified, but can under certain conditions result in more rapid aging as normally expected. The sensors should be treated as systems with a certain life span, which is heavily dependent on the operating regime as shown in the following.

3.1. Ageing due to long-term use and overloads

The pressure sensors are not aging only due to the harsh conditions in which they operate (moisture, temperature), but also due to wear-out effects (the aging of materials due to pressure loads). The situation is similar to that of mechanical parts / systems. Our previous study [16] showed how the offset and the sensitivity characteristics change due to the accelerated ageing with various pressure loads. This article interprets experimental results obtained for the 100-mbar sensors (such as in Fig. 1(a)) in terms of proactive maintenance. The measurements after the loading cycles have shown that at the pressure loads up to full-scale (FS) do not change V_{off} and S critically, and that the scattering of the sensors` readings is acceptably small, while in the case of long-term overloading the drift and the scattering of the readings are much larger. The average changes in the offset voltage of the sensors after accelerated aging (1 million cycles with the pressure loads of 1×FS and 3×FS) measured at 25 °C and 50% RH are presented in Fig. 3. The changes in the sensitivity are shown in Fig. 4.

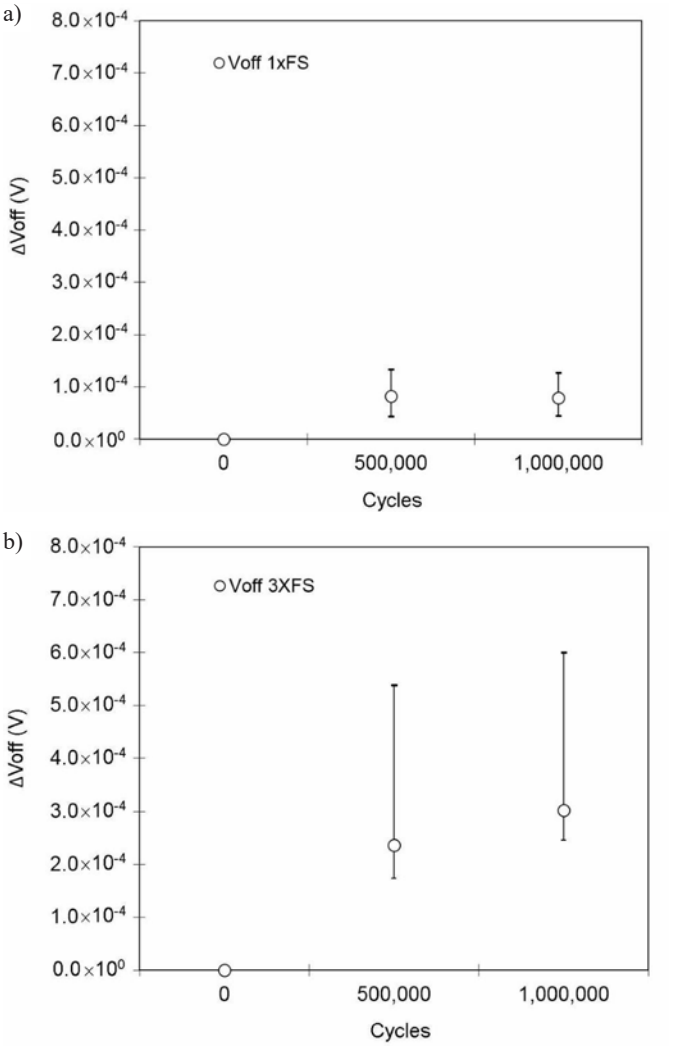


Fig. 3. Results of ageing tests carried out for the series of 10 sensor samples: (a) absolute value of the changes in the offset voltage ($\Delta V_{off} = V_{off} - V_{off0}$) subjected to the pressure cycles 1 × FS, (b) ΔV_{off} in the case of overload cycles (3 × FS)

The above results confirm that the sensors under different loading regimes are aging at different rates. Individual devices may remain good even longer than specified in the producers` specifications, while others are more likely to fall outside the specified accuracy.

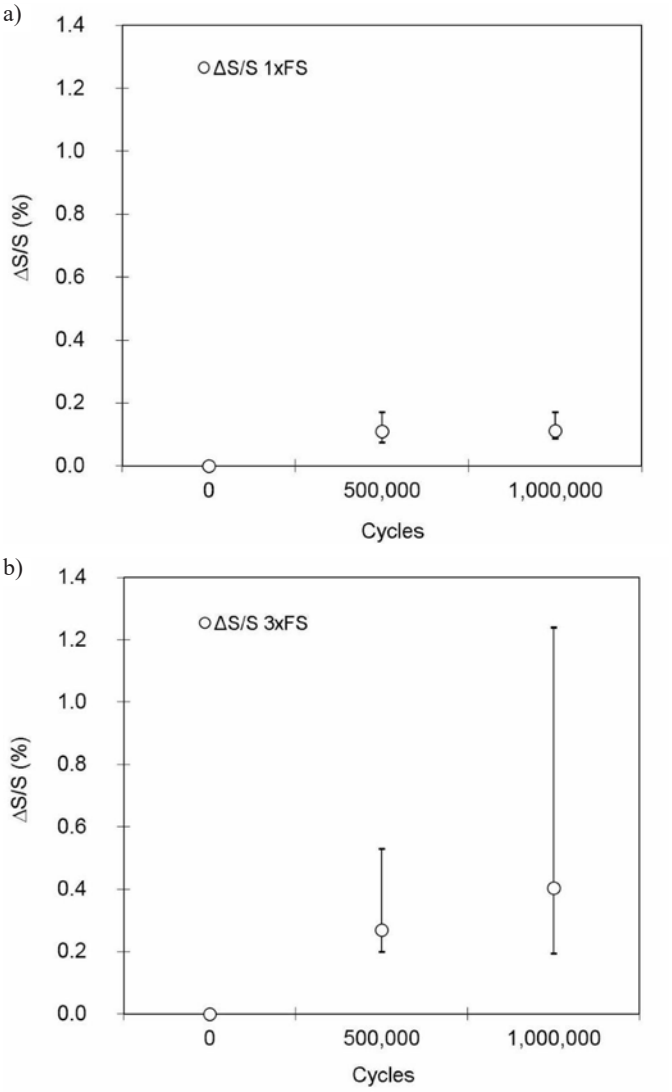


Fig. 4. Absolute value of the relative changes in the sensitivity ($\Delta S/S_0$): (a) sensors subjected to the pressure cycles 1 × FS, (b) sensors subjected to overload cycles 3 × FS

Through a series of similar experiments we assessed the influence of the pressure load intensity on the sensors lifetime. The sensors were exposed to the cycles of the full-scale pressure loads and the overloads that they should undergo according to specifications. The changes in the offset voltage and the sensitivity of a series of 100-mbar pressure sensors after accelerated aging (5000 pressure cycles 1×FS, 3×FS and 5×FS in the air at 25°C and 50% RH) are presented in Fig. 5. The sensors which are more often overloaded can have larger drifts from the calibrated response and reasonable extensive scattering of the measured results.

3.2. Ageing due to the environment/media

Because of good chemical stability of LTCC materials the LTCC-based sensors can be used for wet-wet applications in different media/environments. However, we should be aware of different ageing effects of the measured media and the harsh environmental conditions. In some media/conditions the functional thick-film structure is aging faster than in the dry air or some inert fluids. Glass-containing ceramics are susceptible to stress corrosion in the presence of humidity so that water can be a critical media. Evaluation of such sensors in the water [17] revealed a slightly worse long-term stability and the reduced burst pressure.

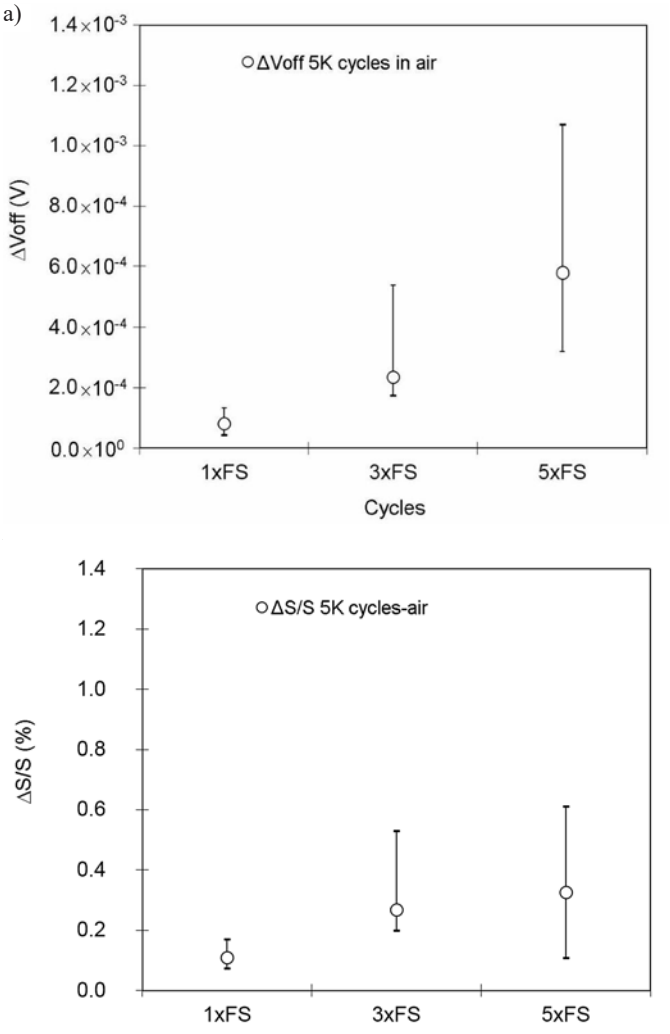


Fig. 5. Measurements of the sensors exposed to the long-term overloading (5000 pressure cycles) with the different load intensities: (a) The change in the offset voltage and (b) The relative change in the sensitivity

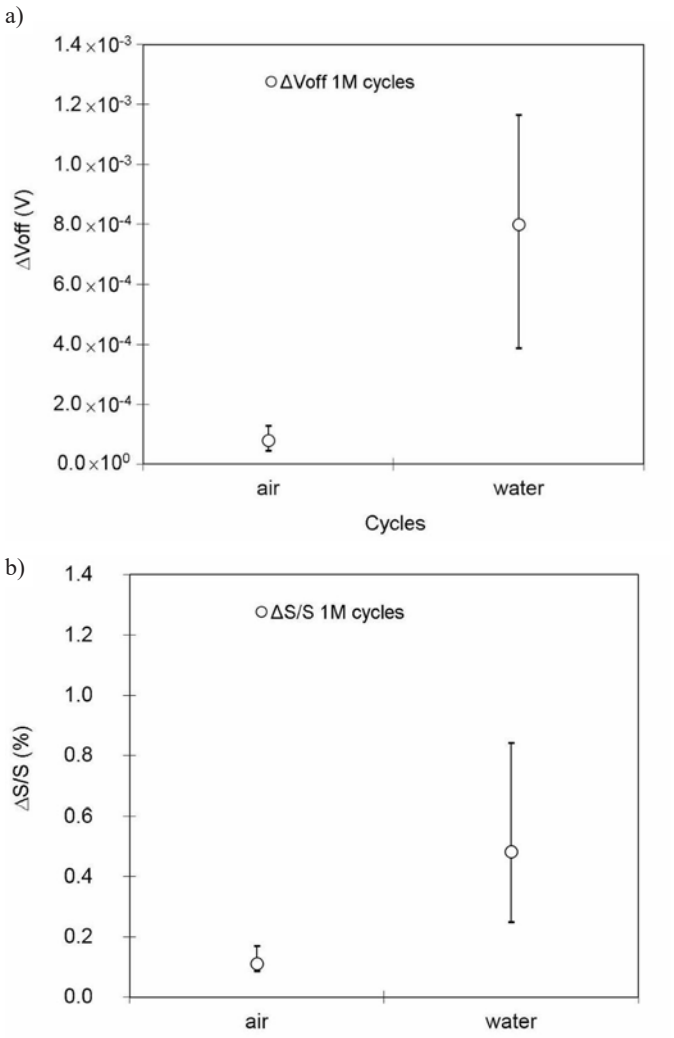


Fig. 6. Characteristics of 100-mbar LTCC sensors after 1 million FS pressure cycles in the air and in water. (a) Changes in the offset voltage, (b) The average relative change in sensitivity.

Comparison of the changes in the offset voltage and the sensitivity of the sensors (shown in Fig. 1(a)) after 1 million pressure cycles in air and in the water are presented in Fig. 6. Notice that the sensors used in this study were not protected against humid environments, and consequently they were disconnected from the power supply during the ageing cycles in the water. Before measurements they were dried in the temperature chamber and stabilized for 24 hours at 50°C.

As evident from Fig. 6, the changes in the sensitivity of the sensor after the loading in the water is significantly greater than after the pressure cycles in the air. Most of the sensors exposed to the pressure loads in the water felt out of the specified tolerances. This implies the reduced lifetime of LTCC sensors in the water. For example, for the required long-term stability of the sensitivity $\Delta S/S_0 < 0.25\%$, the sensor that otherwise may normally withstand one million cycles in the air will fail at the same load in the water.

Taking into account these findings when designing new applications, a sensor can be optimized so that the operation in the water does not critically degrade its reliability. Still, it is important to be aware that the drift of the readings of the sensor operating in the water (due to normal aging and frequent overloads) can be significantly higher than expected for the loads of lower intensity and/or for the operation in the air.

3.3. Ageing due to intrinsic irregularities

As showed in [16] the stability of the sensors' offset voltage and the sensitivity are related to the low-frequency noise of the output signal. A lower noise level corresponds to a better long-term stability while the sensors with higher noise of the output voltage may have poor long term stability.

It may happen that same sensors have small intrinsic irregularities/defects in the material structure which are not detected by typical output tests in serial production. Such sensors are still operating within the specifications, however, experiments showed that irregularities in the thick-film structure may affect their long-term stability and shorten the sensors lifetime.

The low frequency noise of the sensors with small irregularities is typically higher than the noise of the perfect sensors without any irregularity. The measurements of the sensors with small irregularities in the thick-film resistor structure showed for about 30% higher low-frequency noise. As long as the sensor resolution (which is reduced due to the increased signal noise) remains within the range required by application, the higher noise is not critical.

The experiments carried out through the accelerated ageing of a series of 100-mbar LTCC-based sensors, which were subjected to the series of pressure cycles showed that the long-term loading does not cause significant changes in the sensors' signal noise. The situation is evident from Fig. 7 which shows the intrinsic low-frequency noise

of the output signal measured before the ageing cycles and the resulting changes in the offset voltage and the sensitivity after the pressure cycles of intensity full scale range (FS) and the overload pressure of intensity of 3 times FS (3×FS).

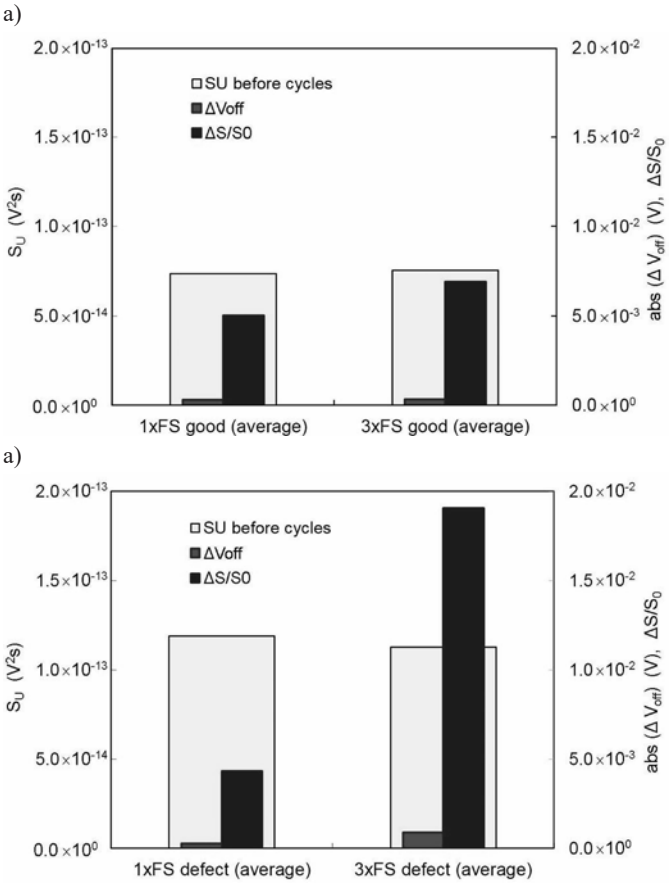


Fig. 7. Average value of the low-frequency noise of the output signal (S_u) measured before the pressure cycles and the changes in the offset voltage and the sensitivity after pressure cycles of intensity 1×FS and 3×FS: (a) the sensors without any defect, (b) the sensors with the defects

Fig. 7(a) shows that the frequently overloading of the sensors with the relatively low noise (< $1e^{-13}$ V^2s) may lead to (in most cases) not critical changes in their sensitivity. The experiments also revealed that the sensors with the intrinsic irregularities have the low-frequency noise > $1e^{-13}$ V^2s (Fig. 7(b)). Measurements of those sensors subjected to overload cycles showed significant drift of their characteristics. As can be seen from Fig. 7(a), the sensors which are free of a defect normally undergo such pressures.

Consequently, the sensors with a higher noise should not be overloaded because operation in such conditions critically impacts their life span. Any irregularities in the structure of the sensor may prove to be hazardous because they can lead into critical defects and the sensor falls out of tolerance. The experiments showed that the pressure cycles of the sensors with small irregularities critically degrade only those sensors which have been overloaded. This suggests that the noise measurements can be used as a successful pre-screening test for a quick assessment of the sensors` long-term stability.

4. Discussion

Based on the above experimental results the following measures have been identified for preventive and predictive maintenance strategies that support the target proactive maintenance.

As regards the preventive maintenance, cyclical maintenance actions should be planned, such as checking the sensor offset after a certain time of operation or a specified number of cyclic loadings. Such subsequent assessment of the sensor´s post occupancy performance makes it possible to examine the interventions planned at the design phase in the light of its actual in-use needs. The subsequent cyclic inspection of sensors health can also include the measurement of the noise of the output signal. As follows from the results presented in 3.3, even a single noise measurement before using the sensor in can be a good indicator of how often the sensor needs to be checked. In this way the offset of the sensors with the high noise ratio of its output signal should be checked in shorter time periods.

The predictive maintenance strategies include the planning of cyclic inspections. In this context regular inspections are carried out to monitor the performance of the design solutions over time and their performance after any intervention. In the above presented case of the LTCC sensors further maintenance actions can be planned after inspection of the sensors offset depending on its degradation level. The larger changes in the offset value mean that the sensor will be out of the required tolerances in a shorter time and need to be regularly inspected after a shorter period. Furthermore, if in addition to the mentioned cyclical inspections, we also collect data giving information about the sensor operating mode (e.g. the frequency of overloads), we can further improve the estimation of the time until the next inspection. As shown in 3.1, 3.2 and 3.3 frequent overloading speeds up the aging of the sensor in all cases.

5. Conclusion

The aging of sensors depends on the measured medium/surroundings as well as on the regime of operation. From the proactive maintenance strategy point of view, we studied the aging of LTCC-based pressure sensors under different operating conditions. In this paper, we show how the sensor response changes due to the long-term operation in the normal conditions and how it's "expected" ageing is affected by the frequent overloading in different environments (i.e., in the air and in the water). We discuss how the aging of LTCC pressure sensors is affected by the long-term operation in normal conditions and under specific conditions with frequent overloading in the air and in the water. In practice, the sensors can withstand a number of overloads (under burst pressure), but at the same time they are aging faster than expected which should be considered in system maintenance. The frequency and intensity of overloads can result in larger drifts from the calibrated response and larger scattering of the measured results. In the water, the overloads are even more critical. Frequent overloading reduces the lifetime of the LTCC sensors more than overloading in the air. In addition, possible intrinsic defects non-detected by the output tests in serial production may also shorten the sensor lifetime under the operation with frequent overloads. Non-negligible effects of aging should be taken into account when selecting the components for proactive maintenance. The proposed measures can be considered also in other sensors/technologies and serve in general as guidelines for implementing maintenance strategies.

Acknowledgement

This work was partially funded from the Slovenian Research Agency [research core funding No. P2-0098].

References

1. Beigl M, Krohn A, Zimmer T, Decker C. Typical Sensors needed in Ubiquitous and Pervasive Computing. Proceedings of the First International Workshop on Networked Sensing Systems (INSS '04), 2004; 153-158.
2. Derler P, Lee EA, Sangiovanni Vincentelli A. Modeling Cyber-Physical Systems. Proceedings of the IEEE 2012, 100(1): 13-28, https://doi.org/10.1109/JPROC.2011.2160929.
3. Fitch J. C. Proactive Maintenance can Yield More than a 10-Fold Savings Over Conventional Predictive/Preventive Maintenance Programs. https://filtagreen-global.com/blog/proactive-maintenance/, (assessed 24. 11.2017).
4. Fitch E C. Proactive Maintenance for Mechanical Systems. Elsevier Science Publishers Ltd, England, 1992.
5. Gerdes M, Galar D, Scholz D. Decision trees and the effects of feature extraction parameters for robust sensor network design. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (1): 31-42, http://dx.doi.org/10.17531/ein.2017.1.5.
6. Gongora-Rubio M R et al., Overview of low temperature co-fired ceramics tape technology for meso-system technology (MsST). Sensors and Actuators A: Physical 2001; 89: 222-241, https://doi.org/10.1016/S0924-4247(00)00554-9.
7. Jakliński P. Analysis of the dual control system operation during failure conditions. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15 (3): 266-272.
8. Jurków D et al., Overview on Low Temperature Co-Fired Ceramic Sensors. Sensors and Actuators A: Phys. 2015; 233: 125-146, https://doi.org/10.1016/j.sna.2015.05.023.
9. Lee EA. The Past, Present and Future of Cyber-Physical Systems: A Focus on Models. Sensors 2015, 15: 4837-4869, https://doi.org/10.3390/s150304837.
10. Lee EA. Fundamental Limits of Cyber-Physical Systems Modeling. ACM Transactions on Cyber-Physical Systems- Inaugural Issue, 2017, 1(1): 3:1-3:28, https://dl.acm.org/citation.cfm?id=3015145&picked=prox&CFID=835065191&CFTOKEN=75865155.
11. Mantis project. http://www.mantis-project.eu/ (accessed 23. 3. 2017).
12. Tolman E. Identifying Pressure Sensor Problems. 2012; http://www.flowcontrolnetwork.com/identifying-pressure-sensor-problems/.
13. Sugier J, Anders GJ. Modelling and evaluation of deterioration process with maintenance activities. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15 (4): 305-311.
14. Swanson L. Linking maintenance strategies to performance. International Journal of Production Economics 2001; 70(3): 237-244, https://doi.org/10.1016/S0925-5273(00)00067-0.
15. Wilson, J S. Sensor Technology Handbook. Amsterdam: Newnes, 2005.
16. Zarnik M S, Sedlakova V, Belavic D, Sikula J, Majzner J, Sedlak P. Estimation of the long-term stability of piezoresistive LTCC pressure sensors by means of low-frequency noise measurements. Sensors and Actuators A: Physical. 2013; 199, 334- 343, https://doi.org/10.1016/j.sna.2013.05.030.
17. Zarnik M S, Belavic D. Study of LTCC-based pressure sensors in water. Sensors and Actuators A: Physical 2014; 220: 45-52, https://doi.org/10.1016/j.sna.2014.09.009.

Marina SANTO ZARNIK
Franc NOVAK
Gregor PAPA
Jožef Stefan Institute
Jamova ulica 39, 1000 Ljubljana, Slovenia

E-mails: marina.santo@ijs.si, franc.novak@ijs.si, gregor.papa@ijs.si

Article citation info:
PENG R. Reliability of interdependent networks with cascading failures. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 273–277, http://dx.doi.org/10.17531/ein.2018.2.13.

Rui PENG

RELIABILITY OF INTERDEPENDENT NETWORKS WITH CASCADING FAILURES

NIEZAWODNOŚĆ WSPÓŁZALEŻNYCH SIECI Z USZKODZENIAMI KASKADOWYMI

The reliability of network systems of various structures has been studied by many researchers. However, most of the works just consider the reliability of a single network system. In practice, different networks may be interdependent such that the failure in one network may result in the failure in another network. The cascading failures have been shown to be catastrophic by some researchers. However, the quantitative evaluation for the reliability of interdependent networks has not been proposed. In this paper, a multi-valued decision diagram based approach is presented to evaluate the reliability of interdependent networks. Illustrative examples are proposed to demonstrate the application of the framework.

Keywords: *reliability, networks, cascading failure, interdependency, multi-valued decision diagram.*

Niezawodność systemów sieciowych o różnych strukturach stanowi przedmiot licznych badań. Jednak większość prac dotyczy tylko niezawodności pojedynczych systemów sieciowych. W praktyce, różne sieci mogą działać współzależnie, tak iż awaria jednej może powodować awarię innej sieci. Niektóre badania pokazują, że uszkodzenia kaskadowe są uszkodzeniami katastroficznymi. Nie zaproponowano jednak dotąd ilościowej oceny niezawodności współzależnych sieci. W niniejszym artykule przedstawiono podejście oparte na koncepcji wielowartościowego diagramu decyzyjnego, które pozwala na ocenę niezawodności wzajemnie zależnych sieci. Przedstawiono przykłady ilustrujące zastosowanie proponowanego paradygmatu.

Słowa kluczowe: *niezawodność, sieci, uszkodzenia kaskadowe, współzależność, wielowartościowy diagram decyzyjny.*

1. Introduction

Researchers have studied the reliability of networks for long [8,9,21]. Typically, they have modeled the reliability of networks with different structures, and have considered different factors, such as common cause failure [1, 7, 25]. [4] studied the influence of cascading failures on the reliability of networks. [6] studied the reliability of networks with multiple terminals using a binary decision diagram (BDD) technique. [22] also used BDD to study the reliability of networks. [12] studied the opportunistic routing for wireless ad hoc and sensor networks. [25] studied the reliability of complex networks with particle swarm optimization approach. [26] studied the optimal link state routing in mobile ad hoc networks. [11] studied the lifetime optimization for a heterogeneous wireless sensor network. [5] studied the reliability of a smart grid network systems considering direct cyber-power interdependency. [19] studied the reliability improvement of a radio electrical distribution network by optimal planning of energy storage systems. [3] studied the reliability enhancement of a wi-fi network. [13] presented the concept of a multi-phase network system to consider dynamic characteristics of networks, and analyzed its reliability. [23] studied the reliability of a cubic network system. [18] presents the method for determining the reliability of a network whose elements (links and nodes) are imperfect (can fail) and repairable. However, most of these works are restricted to the study of a single system.

In practice, the failure of different networks may be interdependent [20,14,10]. Say, the failure in a subway system may increase the load of the bus transportation system, and increases the risk of traffic congestion. Another example is the interdependence between power systems and the control systems. As pointed out in [2], the cascading failure between the power systems and the internet network caused a blackout that affected much of Italy in September 2003. In [2], the effect of removing a proportion of nodes in one network is studied. However, the quantitative evaluation of reliability of interdependent networks is not provided. In this work, a multi-valued decision diagram based approach is adopted to evaluate the reliability of interdependent networks with cascading failures.

Section 2 describes the failure mechanism of the interdependent systems. Section 3 provides the multi-valued decision diagram based approach. Section 4 provides the numerical example. Section 5 concludes.

2. System description

Consider a system consisting of multiple networks, where the failure of some node in a network may cause one or more nodes in another network in fail. Each node in each network has an internal failure rate, and the nodes in each network have known connections with each other. Once a node fails, either due to internal failure or cascading failure, the node and its connections with other nodes are removed from the network it belongs to. After the removal, if any cluster of connected nodes in a network is smaller than a prefixed number, then the cluster will fail. A special case is where a node fails if it is not connected to any other nodes. This kind of cascading failure may cause catastrophic effects, as the failure of a node in one network may result in several nodes in other networks to fail, which may again cause more nodes in the original network to fail. In [2], an illustrative system is proposed, as shown in Fig 1. There are two networks, A and B. Both of them contain six nodes, and the connections of the nodes are shown using the arcs in Fig. 1 (a). Any node will fail in case it is not connected with

any other node in the network. In case A_i or B_i fails (i=1,…,6), B_i or A_i will fail. Therefore, if A_5 fails, then the system will be as shown in Fig. 1 (b) since B_5 will fail and the connections of A_5 and B_5 with other nodes should be removed. Furthermore, as A_4 and A_6 are isolated, they will fail and cause B_4 and B_6 to fail. Afterwards, B_3 becomes isolated, and B_3 and A_3 will fail. Then, the system will be as shown in Fig. 1 (c). That is, the failure of A_5 has caused cascading failures of A_3, A_4, A_6, B_3, B_4, B_5, and B_6. In this paper, the reliability of the interdependent networks is defined as the probability that each network still has some working nodes after a fixed period of time.

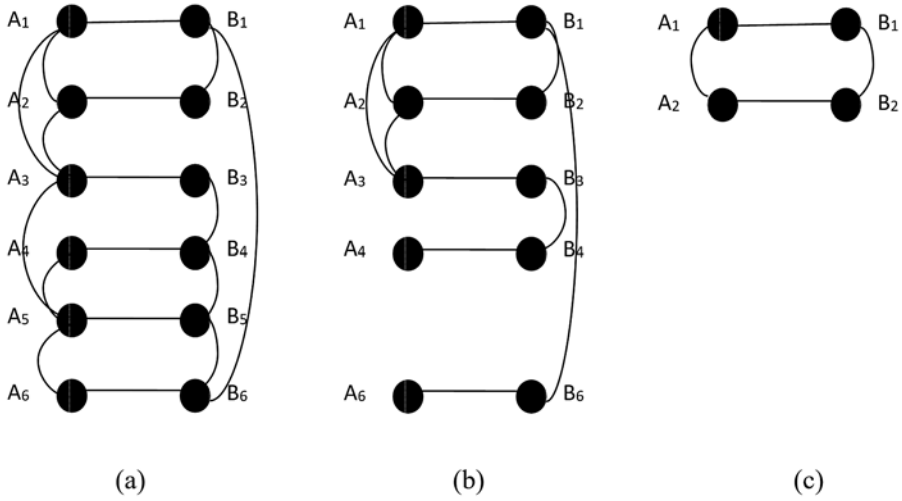


Fig. 1 An illustrative system

3. The model

Multi-valued decision diagram (MDD) has been frequently adopted to evaluate the reliability of systems with dependent failures [16, 17]. However, to adapt to our situation, the MDD used is somewhat different as in most papers. In most papers using MDD, each node in the MDD corresponds to a system element, each branch corresponds to a state of the element, and therefore each path leading to system success represents the set of elements that have failed and the set of elements that have not failed [15, 28]. In our case, if the traditional MDD is used, for each path representing system success, one still needs to enumerate all the possible sequence of the system failures. To avoid enumerate the sequence of failures, similar as in [27], the nodes of our MDD directly represent the failure sequence, and each path leading to system success represents the sequence of failures that have happened. The procedures of evaluating the system reliability with MDD are as follows:

1) Construct the MDD representing the first event, which can be the failure of any node in any network, or no failure happening at all. The terminal for each branch is the set of nodes that have failed in all the networks, considering both internal failures and cascading failures.
2) For the branch representing "no failure" or "no more failure", the terminal for the branch is set to "0" representing system success. For any other branch, if it contains terminal representing that the system is still reliable, the branch needs further branching. The further branches represent all the possible scenarios for the following event, which can be the failure of any remaining node, or no more failure. For any branching indicating system failure, the terminal is set to be "1".
3) Continue step 2 until all the terminals become "0" and"1"..
4) Sum up the probabilities for the paths leading to "0", which is the system reliability.

4. Illustrative example

Consider the illustrative system in Fig. 1 (a), and assume that the system is reliable as long as at least two connected nodes are working in each network. According to the procedures, the MDD for the illustrative system shown in Fig. 1 (a) can be constructed. In order to make the MDD more concise, we do not show the branches directly leading to "0" and "1". The MDD for the system is as shown in Fig. 2.

From the MDD, the scenarios that lead to system success can be summarized below:

Scenario 1: No failure.

Scenario 2: A_1 or B_1 fails, leading to the failure of A_1, A_2, B_1, B_2, then no more failure.

Scenario 3: A_1 or B_1 fails, then A_3 or B_3 fails, leading to the failure of A_1-A_3 and

B_1-B_3, then no more failure.

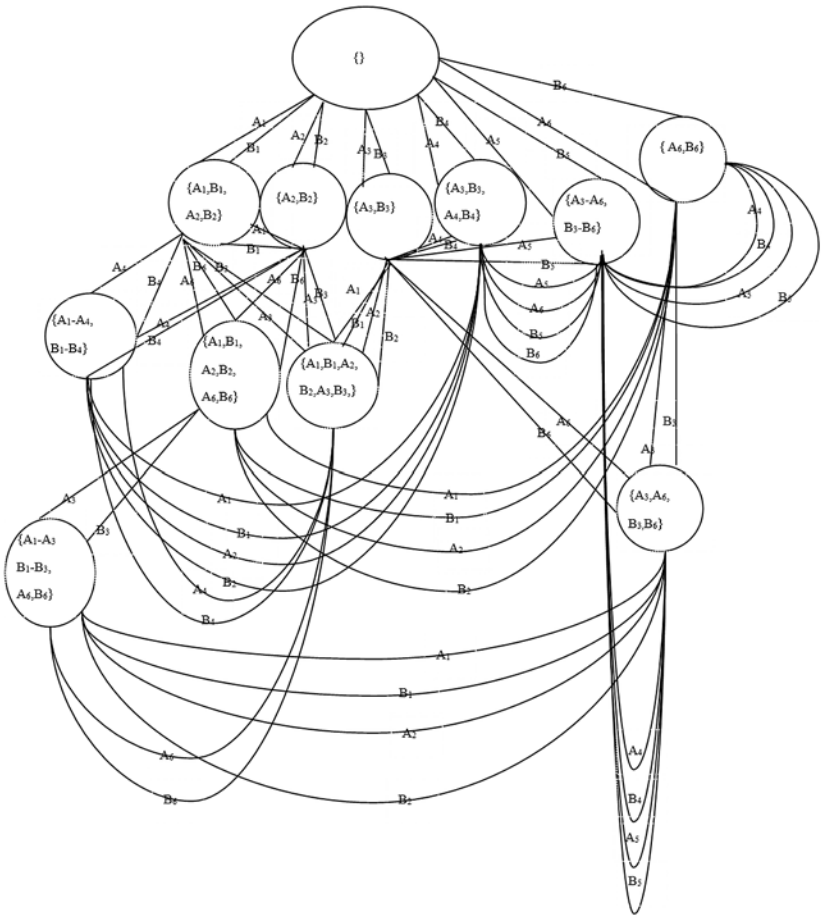


Fig. 2. MDD for the illustrative system in Fig. 1. (a)

Scenario 4: A_1 or B_1 fails, then A_3 or B_3 fails, then A_4 or B_4 fails, leading to the failure of A_1-A_4,B_1-B_4,,then no more failure.

Scenario 5: A_1 or B_1 fails, then A_3or B_3 fails, then A_6 or B_6 fails, leading to the failure of A_1-A_3,B_1-B_3, A_6,B_6,,then no more failure.

Scenario 6: A_1 or B_1 fails, then A_4, or B_4 fails, leading to the failure of A_1-A_4 and B_1-B_4, then no more failure.

Scenario 7: A_1 or B_1 fails, then A_6 or B_6 fails, leading to the failure of A_1,B_1,A_2,B_2,A_6,B_6,, then no more failure.

Scenario 8: A_1 or B_1 fails, then A_6 or B_6 fails, then A_3 or B_3 fails, leading to the failure of A_1- A_3, B_1- B_3,A_6,B_6,, then no more failure.

Scenario 9: A_2 or B_2 fails, leading to the failure of A_2, B_2, then no more failure.

Scenario 10: A_2 or B_2 fails , then A_1 or B_1 fails, leading to the failure of A_1,A_2,B_1,B_2, then no more failure.

Scenario 11: A_2 or B_2 fails, then A_1 or B_1 fails, then A_3,or B_3 fails, leading to the failure of A_1-A_3 and B_1-B_3, then no more failure.

Scenario 12: A_2 or B_2 fails, then A_1 or B_1 fails, then A_3or B_3 fails, then A_4 or B_4 fails, leading to the failure of A_1-A_4 and B_1-B_4, then no more failure.

Scenario 13: A_2 or B_2 fails , then A_1 or B_1 fails, then A_3,or B_3 fails, then A_6 or B_6 fails, leading to the failure of A_1-A_3,B_1-B_3, A_6,B_6,,then no more failure.

Scenario 14: A_2 or B_2 fails, then A_1 or B_1 fails, then A_4, or B_4 fails, leading to the failure of A_1-A_4 and B_1-B_4, then no more failure.

Scenario 15: A_2 or B_2 fails , then A_1 or B_1 fails, then A_6 or B_6 fails, leading to the failure of A_1,B_1,A_2,B_2,A_6,B_6,, then no more failure.

Scenario 16: A_2 or B_2 fails , then A_1 or B_1 fails, then A_6 or B_6 fails, then A_3 or B_3 fails, leading to the failure of A_1- A_3, B_1- B_3,A_6,B_6,, then no more failure.

Scenario 17: A_2 or B_2 fails, then A_3 or B_3 fails, leading to the failure of A_1- A_3, B_1- B_3, then no more failure.

Scenario 18: A_2 or B_2 fails, then A_3 or B_3 fails, then A_4 or B_4 fails, leading to the failure of A_1- A_4, B_1- B_4, then no more failure.

Scenario 19: A_2 or B_2 fails, then A_3 or B_3 fails, then A_6 or B_6 fails, leading to the failure of A_1-A_3, B_1- B_3, A_6,B_6,, then no more failure.

Scenario 20: A_2 or B_2 fails, then A_4 or B_4 fails, leading to the failure of A_1- A_4, B_1- B_4, then no more failure.

Scenario 21: A_2 or B_2 fails, then A_6 or B_6 fails, leading to the failure of A_1,B_1,A_2,B_2,A_6,B_6,, then no more failure.

Scenario 22: A_2 or B_2 fails, then A_6 or B_6 fails, then A_3 or B_3 fails, leading to the failure of A_1- A_3, B_1- B_3,A_6,B_6, then no more failure.

Scenario 23: A_3 or B_3 fails, leading to the failure of A_3, B_3, then no more failure.

Scenario 24: A_3 or B_3 fails, then A_1,A_2,B_1, or B_2 fails, leading to the failure of A_1-A_3 and B_1-B_3, then no more failure.

Scenario 25: A_3 or B_3 fails, then A_1,A_2,B_1, or B_2 fails, then A_4 or B_4 fails, leading to the failure of A_1-A_4 and B_1-B_4, then no more failure.

Scenario 26: A_3 or B_3 fails, then A_1,A_2,B_1, or B_2 fails, then A_6 or B_6 fails, leading to the failure of A_1-A_3 and B_1-B_3, A_6,B_6, then no more failure.

Scenario 27: A_3 or B_3 fails, then A_4 or B_4 fails, leading to the failure of A_3,A_4,B_3,B_4, then no more failure.

Scenario 28: A_3 or B_3 fails, then A_4 or B_4 fails, then A_1,A_2,B_1,B_2 fails, leading to the failure of A_1-A_4,B_1-B_4, then no more failure.

Scenario 29: A_3 or B_3 fails, then A_4 or B_4 fails, then A_5,A_6,B_5,B_6 fails, leading to the failure of A_3-A_6,B_3-B_6, then no more failure.

Scenario 30: A_3 or B_3 fails, then A_5 or B_5 fails, leading to the failure of A_3-A_6,B_3-B_6, then no more failure.

Scenario 31: A_3 or B_3 fails, then A_6 or B_6 fails, leading to the failure of A_3,A_6,B_3,B_6, then no more failure.

Scenario 32: A_3 or B_3 fails, then A_6 or B_6 fails, then A_1,A_2, B_1 or B_2 fails, leading to the failure of A_1- A_3, B_1- B_3,A_6,B_6,then no more failure.

Scenario 33: A_3 or B_3 fails, then A_6 or B_6 fails, then A_4,A_5, B_4 or B_5 fails, leading to the failure of A_3- A_6, B_3- B_6,then no more failure.

Scenario 34: A_4or B_4 fails, leading to the failure of A_3,A_4,B_3,B_4, then no more failure.

Scenario 35: A_4or A_4 fails, then A_1,A_2,B_1,B_2 fails, leading to the failure of A_1-A_4,B_1-B_4, then no more failure.

Scenario 36: A_4or A_4 fails, then A_5,A_6,B_5,B_6 fails, leading to the failure of A_3-A_6,B_3-B_6, then no more failure.

Scenario 37: A_5 or B_5 fails, leading to the failure of A_3-A_6, B_3-B_6, then no more failure.

Scenario 38: A_6 or B_6 fails, leading to the failure of A_6, B_6, then no more failure.

Scenario 39: A_6 or B_6 fails, then A_1,A_2,B_1,B_2 fails, leading to the failure of A_1,B_1,

A_2,B_2,A_6,B_6, then no more failure.

Scenario 40: A_6 or B_6 fails, then A_1,A_2,B_1,B_2 fails, then A_3, B_3 fails, leading to the failure of A_1-A_3,B_1-B_3,A_6,B_6, then no more failure.

Scenario 41 A_6 or B_6 fails, then A_3 or B_3 fails, leading to the failure of A_3,A_6,B_3,B_6, then no more failure.

Scenario 42: A_6 or B_6 fails, then A_3 or B_3 fails, then A_1,A_2, B_1 or B_2 fails, leading to the failure of A_1- A_3, B_1- B_3,A_6,B_6,then no more failure.

Scenario 43: A_6 or B_6 fails, then A_3 or B_3 fails, then A_4,A_5, B_4 or B_5 fails, leading to the failure of A_3- A_6, B_3- B_6,then no more failure.

Scenario 44: A_6 or B_6 fails, then A_4,A_5,B_4,B_5 fails, leading to the failure of A_3-A_6, B_3-B_6, then no more failure.

Note that though the enumeration of all the scenarios seems to be tedious, it is actually done according to a depth-first traversal. For small examples, one can enumerate the scenarios manually, whereas one needs to construct the MDD with computer programming and then sort out all the paths leading to system success through either depth-first traversal or width-first traversal if the system has a larger scale. Indeed, we admit that the system MDD can grow fast when the networks have more nodes, but it is also not supposed to solve the reliability of a complicated system with simple steps. Fortunately, with the advancement of computing technology, such as parallel computing and quantum computing, it is promising for the computer to analyze a MDD with thousands of nodes in seconds.

Assume that the system operation time is T. The failure time of each node observes exponential distribution, with failure rate λ_i for A_i and β_i for B_i. The system reliability can be obtained by summing up the probabilities for all the scenarios leading to system success. Set λ_i=β_i=0.01 for i=1,..,6 and T=20, the system reliability can be calculated to be R= 0.7783. The influence of different nodes on system

Table 1. System reliability when changing failure rate of different nodes

Cases	Benchmark	Change λ_1 or β_1 to 0.002	Change λ_2 or β_2 to 0.002	Change λ_3 or β_3 to 0.002	Change λ_4 or β_4 to 0.002	Change λ_5 or β_5 to 0.002	Change λ_6 or β_6 to 0.002
System Reliability	0.7783	0.7346	0.7455	0.7783	0.7635	0.7187	0.7709

reliability is studied by calculating the system reliability again by changing λ_i and β_i to 0.02 and keeping other parameters unchanged. Table 1 shows the results. It can be seen that increasing the failure rate of node A_3 and B_3 does not have much influence on the system reliability. Actually, when A_3 or B_3 fails, A_1,A_2,A_4,A_5,A_6,B_1,B_2,B_4,B_5,B_6 can still function. Similarly, changing the failure rate of A6 or B6 also has minor effects. Actually, when A_6 or B_6 fails, A_1-A_5 and B_1-B_5 can still function. Increasing the failure rate of A_5 or B_5 has the biggest effect. Actually, when A_5 or B_5 fails, A_3-A_6 and B_3-B_6 will all fail due to cascading effects.

5. Conclusions

This paper proposed a multi-valued decision diagram based approach to evaluate the reliability of interdependent networks. Any node in each network has an intrinsic failure rate, and the failure of it may cause some nodes in other networks to fail. Moreover, a cluster of connected nodes fail as long as its size is smaller than a pre-specified number. A special case is where any node fails as long as it is not connected to any other nodes. The system is considered as reliable as long as it still has some working nodes in each network after a fixed period of time.

In this work, the failure of a node will cause fixed nodes to fail. It would be interesting to consider the case where a node failure may cause a random set of nodes to fail. Another direction is to consider the case where each node is multi-state instead of binary state. In the future, works can be done to calculate the importance measures of different nodes, and investigate the optimal structure of the networks. Besides, for very big networks, directly adopting the procedures may be computational complicated and unnecessary. In the future, works can be done to divide interdependent complicated networks into interdependent clusters, and calculate the reliability of the dependent networks based on the reliability of each cluster and the relationship of different clusters.

Acknowledgement

The research reported here was partially supported by the NSFC under grant numbers 71671016, 71231001, and 71420107023, and the Fundamental Research Fund of Central Universities under the grant number FRF-GF-17-B14.

References

1. Albert R, Jeong H, Barabasi AL. Error and attack tolerance of complex networks. Nature 2000; 406: 378-382, https://doi.org/10.1038/35019019.
2. Buldyrev S, Parshani R, Paul G, Stanley H, Havlin S. Catastrophic cascade of failures in interdependent networks. Nature 2010; 494 (15): 1025-1028, https://doi.org/10.1038/nature08932.
3. Cena G, Scanzio S, Valenzano A. Seamless Link-Level Redundancy to Improve Reliability of Industrial Wi-Fi Networks. IEEE Transactions on Industrial Informatics 2016; 12 (2): 608-620, https://doi.org/10.1109/TII.2016.2522768.
4. Crucitti P, Latora V, Marchiori M. Model for cascading failures in complex networks. Physical Review E 2004; 69(4): 045104, https://doi.org/10.1103/PhysRevE.69.045104.
5. Falahati B, Fu Y, Wu L. Reliability assessment of smart grid considering direct cyber-power interdependencies. IEEE Transactions on Smart Grid 2012; 3 (3): 1515-1524, https://doi.org/10.1109/TSG.2012.2194520.
6. Hardy G, Lucet C, Lininios N. K-terminal network reliability measures with binary decision diagrams. IEEE Transactions on Reliability 2007; 56(3): 506-515, https://doi.org/10.1109/TR.2007.898572.
7. Kuhnle A, Nguyen NP, Dinh TN, Thai MP. Vulnerability of clustering under node failure in complex networks. Social Network Analysis and Mining 2017; 7 (1): 8-23.
8. Levitin G, Gertsbakh I, Shpungin Y. Evaluating the damage associated with intentional supply deprivation in multi-commodity network. Reliability Engineering & System Safety 2013; 119: 11-17, https://doi.org/10.1016/j.ress.2013.05.002.
9. Levitin G, Xing L, Dai Y. Optimal data partitioning in cloud computing system with random server assignment. Future Generation Computer Systems-The International Journal of Escience 2015; 70: 17-25, https://doi.org/10.1016/j.future.2016.12.025.
10. Lin Y, Kang R, Wang Z, Zhao Z, Li D, Havlin S. Robustness of networks with dependency topology. EPL 2017; 118: 36002, https://doi.org/10.1209/0295-5075/118/36002.
11. Lin Y, Zhang J, Chung H, Ip WH, Li Y, Shi Y. An ant colony optimization approach for maximizing the lifetime of heterogeneous wireless sensor networks. IEEE Transactions on Systems Man and Cybernetics Part C-Applications and Reviews 2012; 42(3): 408-420, https://doi.org/10.1109/TSMCC.2011.2129570.
12. Liu H, Zhang B, Mouftah H, Shen X, Ma J. Opportunistic routing for wireless ad hoc and sensor networks: present and future directions. IEEE Communications Magazine 2009; 47(12): 103-109, https://doi.org/10.1109/MCOM.2009.5350376.
13. Lu JM, Innal F, Wu XY, Liu YL, Lundteigen M. Two-terminal reliability analysis for multi-phase communication networks. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (3): 418-427, https://doi.org/10.17531/ein.2016.3.14.
14. Majdandzic A, Braunstein L, Curme C, Vodenska I, Levy-Carcienta S, Stanley H, Havlin S. Multiple tipping points and optimal repairing in interacting networks. Nature Communications 2016; 7: 10850, https://doi.org/10.1038/ncomms10850.
15. Mo Y, Xing L, Zhong F, Zhang Z. Reliability evaluation of network systems with dependent propagated failures using decision diagrams. IEEE Trans. Depend. Secure Comput 2016; 13 (6): 672-683, https://doi.org/10.1109/TDSC.2015.2433254.
16. Mo Y, Xing L, Cui L, Si S. MDD-based performability analysis of multi-state linear consecutive-k-out-of-n: F systems. Reliability Engineering & System Safety 2-17; 166: 124-131, https://doi.org/10.1016/j.ress.2016.08.027.

17. Peng R, Zhai QQ, Xing LD, Yang J. Reliability of demand-based phased-mission systems subject to fault level coverage. Reliability Engineering and System Safety 2014; 121: 18-25, https://doi.org/10.1016/j.ress.2013.07.013.
18. Pilch R. Reliability evaluation of networks with imperfect and repairable links and nodes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (1): 19-25, http://dx.doi.org/10.17531/ein.2017.1.3.
19. Saboori H, Hemmati R, Jirdehi M. Reliability improvement in radial electrical distribution network by optimal planning of energy storage systems. Energy 2015; 93: 2299-2312, https://doi.org/10.1016/j.energy.2015.10.125.
20. Shekhtman L, Danziger M, Havlin S. Recent advances on failure and recovery in networks of networks. Chaos Solutions & Fractals 2016; 90: 28-36, https://doi.org/10.1016/j.chaos.2016.02.002.
21. Tchórzewska-Cieślak B, Pietrucha-Urbanik K, Urbanik M. Analysis of the gas network failure and failure prediction using the Monte Carlo simulation method. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (2): 254-259, http://dx.doi.org/10.17531/ein.2016.2.13.
22. Xing L. An efficient binary-decision-diagram-based approach for network reliability and sensitivity analysis. IEEE Transactions on Systems Man and Cybernetics Part A-Systems and Humans 2008; 38 (1): 105-115, https://doi.org/10.1109/TSMCA.2007.909493.
23. Xu X, Zhou S, Li J. Reliability of complete cubic networks under the condition of g-good-neighbour. Computer Journal 2017; 60 (5): 625-635.
24. Yeh WC, Lin YC, Chung Y, Chih M. A particle swarm optimization approach based on Monte Carlo simulation for solving the complex network reliability problem. IEEE Transactions on Reliability 2010; 59 (1): 212-221, https://doi.org/10.1109/TR.2009.2035796.
25. Yeh W C. A squeezed artificial neural network for the symbolic network reliability functions of binary-state networks. IEEE Transactions on eural Networks and Learning Systems 2017; 28 (11): 2822-2825, https://doi.org/10.1109/TNNLS.2016.2598562.
26. Yi J, Adnane A, David S, Parrein B. Multipath optimized link state routing for mobile ad hoc networks. Ad Hoc Networks 2011; 9(1): 28-47, https://doi.org/10.1016/j.adhoc.2010.04.007.
27. Zhai QQ, Peng R, Xing LD, Yang J. Reliability of demand-based warm standby systems subject to fault level coverage. Applied Stochastic Models in Business and Industry 2015; 31 (3): 380-393, https://doi.org/10.1002/asmb.2010.
28. Zhang S, Sun S, Si S, Wang P. A decision diagram based reliability evaluation method for multiple phased-mission systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (3): 485-492, http://dx.doi.org/10.17531/ein.2017.3.20.

Rui PENG
RM1119, Donlinks School of Economics & Management
University of Science & Technology Beijing
Xueyuan Rd 30, Haidian district, Beijing, China

Email: pengrui1988@ustb.edu.cn

Article citation info:
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.

Stanisław MŁYNARSKI
Robert PILCH
Maksymilian SMOLNIK
Jan SZYBKA

METHODOLOGY OF NETWORK SYSTEMS RELIABILITY ASSESSMENT ON THE EXAMPLE OF URBAN TRANSPORT

METODYKA SZACOWANIA NIEZAWODNOŚCI UKŁADÓW SIECIOWYCH NA PRZYKŁADZIE KOMUNIKACJI MIEJSKIEJ*

Apart from reliability evaluation, the methodology of network systems reliability assessment presented in the article enables the design of modernisation of such systems targeted mainly at ensuring their required reliability. In practice the methodology can be applied for various network systems, e.g. computer, power, gas, water distribution, telecommunications and transport networks. A reliability analysis of a transport network in urban public transport is presented. Calculations were performed for selected criteria of network availability which actually conditions the quality of transport services provided. The basic calculation tool used was the factoring algorithm that enabled the assessment of the impact of individual connections failure (in particular those caused by physical factors) on the reliability of the whole network. The feasibility of modernisation of the network analysed is discussed and the results are presented in diagrams.

***Keywords**: transport network, reliability, factoring algorithm.*

W artykule zaprezentowano opracowaną metodykę szacowania niezawodności układów sieciowych. Rozwiązanie to umożliwia dokonywanie oceny niezawodności oraz projektowanie modernizacji rozpatrywanej sieci przede wszystkim w aspekcie zapewnienia jej wymaganej niezawodności. Praktyczne wykorzystanie omawianej metodyki może mieć miejsce w odniesieniu do różnych układów sieciowych, np. sieci komputerowych, energetycznych, gazowych, wodociągowych, telekomunikacyjnych i transportowych. W artykule przedstawiono analizę niezawodności sieci komunikacyjnej w miejskim transporcie zbiorowym. Obliczenia przeprowadzono dla wybranych kryteriów zdatności sieci, które praktycznie warunkują jakość świadczonych usług transportowych. Podstawowe narzędzie obliczeniowe stanowił algorytm faktoryzacji, który umożliwia ocenę wpływu uszkodzeń poszczególnych połączeń (spowodowanych w szczególności czynnikami fizycznymi) na niezawodność całej sieci. W opracowaniu uwzględniono możliwość modernizacji analizowanej sieci, a uzyskane wyniki przedstawiono na wykresach.

***Słowa kluczowe**: sieć komunikacyjna, niezawodność, algorytm faktoryzacji.*

1. Introduction

In general terms, the paper deals with issues of network reliability assessment, which due to the specificity of such networks cannot be done with traditional methods used in the theory of reliability. In the paper various types of systems are characterised, a methodology of reliability calculations procedure for network systems is presented. Moreover, using a case of urban transport the usefulness of the factorisation method applied is proved.

The activities of municipal transport companies are closely connected with the management of fixed assets whose components include transportation means and infrastructure, i.e. transport routes and networks. An important element of transportation means use is constituted by the maintenance-repair management, which is an indispensable condition for ensuring the continuity of operation and utilisation of the assets potential [2, 6, 27].

This management covers all the undertakings aimed at continuing the serviceability of transport networks. These undertakings involve the necessity of temporary closure of some lines and negatively affect the effects of company's operation. The aim of companies is to design the structure of connections in such a way that in the case of emergency or pre-planned repair of traction lines the transport system should guarantee users reaching their destination point, (cf. [10, 14, 31]).

Meeting the increasing demands in this area makes it necessary to look for methods of modelling and evaluating transportation systems taking into account emergencies and economic criteria. To achieve this aim the probability calculus is used (cf. [11, 27, 28, 33]). For instance, the probabilistic index of trips irregularity defined as the difference (in minutes) between the real and scheduled time in the conditions of ensuring stable and regular traffic is analysed. This index is affected by the transport network shape. Its favourable configuration is significant for accomplishing the transportation task, (cf. [9, 18]).

To solve this task the factoring algorithm proposed in the paper can be applied [15, 24]. It enables determining the reliability for complex network systems used in municipal transport systems and is an alternative to simulation methods applied in such cases, which, however, are less precise [17, 19, 20].

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

2. Characteristics of network systems

Network systems include power supply, IT, water supply, transportation that ensure the operational continuity of a company and product distribution as well as other systems of similar type.

All the above belong to technical infrastructure which covers facilities, transmission networks and related technical objects performing tasks of supply of heat, gas and water, wastewater and waste disposal, transportation, etc. Infrastructure denotes essential facilities and service institutions indispensable for the operation of economy and society. Linear infrastructure in a network system is characterised by the linear connection between the technical means (components) that perform certain functions.

To evaluate network systems block diagram based methods can be used. These can be applied for reliability module structures and extensive transportation systems in urban agglomerations that operate on a large area and are composed of numerous nodes that contain mapped internal subsystems [12].

Network structures are frequently modelled using graphs composed of nodes and sections connecting the nodes. The reliability of structures mapped as graphs can be effectively calculated using factorisation methods [17, 19, 23]. The reliability structures in municipal transport systems can be subject to continuous evolution. This may be because the area covered by the operation may change, or the demand for transportation services may change. The correct identification of real-life structures of technical systems together with the method of calculating reliability indices are of key significance for the proper reliability analysis and assessment of threat to the effective operation of a municipal transport system.

What makes the assessment of network system reliability difficult is the fact that they do not have reliability structure generally considered classic, there may occur interdependencies of failures between parts of the system, and the application of typical reliability indices and characteristics is limited. This is why the reliability assessment of this type of structures can be done using the factorisation algorithm making use of the graph theory and the method of network reduction [4, 23, 24]. The shaping and optimisation of the structure of connections of a network type system as early as the design stage is an important aspect of ensuring its operational reliability [5, 13, 26]. The credibility of network reliability assessment can be improved by taking into account the failures and recoveries in the operational process of all the elements of the network [25].

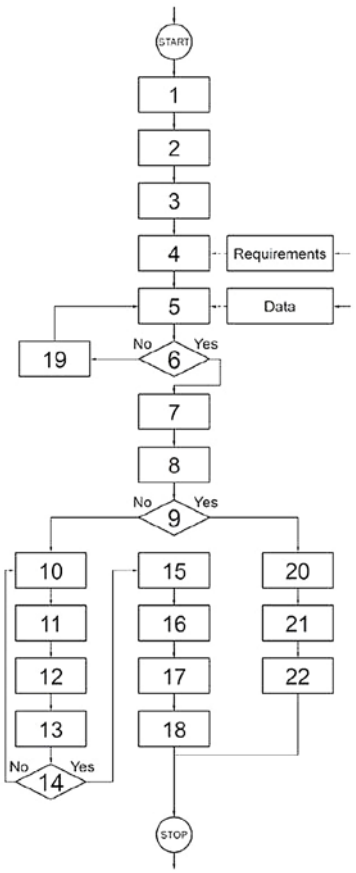
3. Methodology of network systems reliability testing

The assessment of network systems reliability requires the use of mathematical models and methods involved in their application. This activity, however, should be preceded by preliminary works and followed by relevant processing of the obtained results [30]. All the activities that may accompany network system reliability assessment have been systematised in the framework of the network assessment and modernisation method developed. Its scheme is presented in figure 1.

The immediate goal of a network system analysis may be the specification of the indices characterising its reliability, but this should merely support more complex activities of more significant impact. These include, in particular, scheduling of preventive maintenance of the network and preparation of network modernisation (especially by modification of its reliability structure).

Preventive maintenance scheduling is outside the scope of the method proposed, however, it does point out that modernisation projects should be preceded by thorough evaluation of the network system reliability.

As shown in figure 1, reliability assessment requires preliminary steps and precisely formulated requirements as to the correct operation of the network system analysed together with data on its operation and maintenance history so far. When the quality of the available data is not sufficient, equivalent data sources on the operation process can be used, for instance accounting records, records of utilisation of spare parts and materials, etc (cf. [30]).



1. Formulate the aim of analysis.
2. Model the network in question as a system of nodes and links.
3. Name individual nodes and links.
4. Specify the function of the network and criterion of its availability.
5. Verify network reliability data.
6. Is the quality of network reliability data sufficient?
7. Process data on individual nodes and links.
8. Evaluate network reliability according to criterion (cf. 4), using mathematical models.
9. Is network reliability evaluation the sole aim of the analysis (cf. 1)?
10. Identify the "weak links" of the network.
11. Develop network modernisation variant.
12. Adopt expected reliability of network new sections.
13. Assess reliability $R_P(t)$ of modernised network.
14. Is $R_P(t) \geq R_W(t)$?
15. Check the entire solution.
16. Develop relevant technical documentation.
17. Verify documentation accuracy.
18. Approve the entire solution and documentation.
19. Plan and perform relevant operational tests or use data equivalent sources.
20. Process analysis results.
21. Verify analysis accuracy.
22. Save analysis results.

Fig. 1. Procedure of the method of network system reliability assessment and modernisation

The activities and questions shown in the scheme by means of blocks with corresponding numbers (steps 1 to 22) are given in the figure. The presented block diagram describes the procedure in relation to which the reliability analysis of the tram network characterised in Chapter 4 was carried out. The descriptions in the individual blocks shown in figure 1 are general in nature which is a consequence of the universality of the developed method that can be used during various analyses and design works concerning the reliability of network systems.

In the method proposed it is assumed that network will be modernised by introducing changes in its reliability structure until its estimated reliability $R_P(t)$ becomes not less than required $R_W(t)$ – this is done in the iteration loop covering steps from 10 to 14. This method of reliability formation (for technical objects of a different class) was presented in [7, 8], (cf. [29]).

4. An example of performing a reliability analysis of a transport network in urban public transport

The requirements for urban transport services quality are highly dependent on the organisation system, technical means of transport and structure of transport connections network together with the possibility of using it. The quality and reliability of urban public transport are determined by the criteria of high significance for both the user

and operator, who must also take into account factors related to the economy of use.

From among the quality and economy criteria in urban transport those of strategic importance include the infrastructure of connections network and vehicles performing the transportation services. The most significant criteria include [1, 3, 16, 22]:

– route availability resulting from the number and distribution of stops,
– spatial and functional integration with other public transport means (national, regional and local, *Park and Ride, Bike and Ride*, etc.),
– direct connections and functional flexibility understood as the possibility of choice of an alternative connection,
– functional reliability whose criteria include: punctuality, regularity, probability of reaching the end point of route.

The transport network structure determines the possibility of connections between nodes, which in turn affect directly the transportation potential of the entire system and the formation of users' opinions of the quality of the transport system. The term of node in a transport network structure denotes the location of tracks intersections and junctions and stops enabling changing to another means of transport, as well as the start and end stops. Rationally arranged infrastructure in the form of effective location of strategic transport nodes offers the possibility to optimise the connections network and the resulting high reliability of the system. What is important for the effectiveness and quality of the entire system is also the high reliability of transport nodes and sections, with particular attention to nodes of strategic significance.

The transport connections structure of an agglomeration can be presented as a graph that shows the functional relations between the components of the system, and analysed in the reliability framework. It can be interpreted as a connections network in urban rail transport between the infrastructure nodes. The transportation process is performed between the initial nodes (depots) and end nodes (tram reversing loops). The other components are the system nodes as partial execution of tasks (tram stops), while the lines are the rail infrastructure between nodes. In such interpretation the availability or unavailability of a section account for the transport service execution by one or more lines or failure to execute the service by the subordinate (any of them) line from beginning to end. It was assumed that the system availability corresponds to the execution of the service by all lines, while unavailability is interpreted as lack of execution of service even by one line [10, 16].

The application of factorisation algorithm in the network structural analysis for the assessment of reliability of connections in urban public transport is proposed.

The factorisation method is based on the following assumptions [4, 21, 25, 32]:

- analysed network is represented by undirected graph $G=(V, E)$, in which $V=\{v_1, v_2,\ldots, v_n\}$ represents set of vertices (nodes in network) and $E=\{e_1, e_2,\ldots, e_m\}$ represents set of edges (links in network)

- all links e_i in the network can fail statistically independently of each other with known probability

- the measure of network reliability is the probability that all nodes from specified set K are connected and: $2 \le |K| \le |V|$

– - specified set K of nodes determines the measure of network reliability (K – terminal reliability network)

– - the reduction process is based on a well-known principle of contracting and deleting of links which is recursively applied

for all edges e_i graph G, which can be rewritten as:

$$R_s = R(G_K) = R_{e_i} \cdot R\left(G_{K'} * e_i\right) + \left(1 - R_{e_i}\right) \cdot R\left(G_K - e_i\right) \qquad (1)$$

where:

e_i – i-th link in a network represents by graph G

R_{ei} – probability that link e_i is in an operating state

$$G_{K'} * e_i = \left(V - v_k - v_l + v_{kl}, E - e_i\right), \quad v_{kl} = v_k \cup v_l,$$

$$K' = \begin{cases} K & \text{if } v_k \text{ or } v_l \notin K, \\ K - v_k - v_l + v_{kl} & \text{if } v_k \in K \text{ and } v_l \in K, \end{cases}$$

$$G_K - e_i = \left(V, E - e_i\right),$$

– as the effect of reduction process one can obtain the formula which represents analytical form of reliability structure and enables for computation the specified measure of network reliability.

The factorisation method makes it possible to calculate the reliability of network systems, which is not possible with the use of classical methods of reliability evaluation. The effect of these calculations is, of course, obtaining typical reliability characteristics of the entire system, however the path leading to their determination is not within the standard procedure for the classical approach to estimating the reliability of technical objects.

Urban transport systems can be mapped as network structures for which many inputs z and outputs w_k should be distinguished, while in the graph representation we have to do with graphs or networks as shown in figure 1. As the availability condition of such a graph the existence of a connection from each initial node (input)) $z \in \boldsymbol{Z}$ to any final node (output) $w \in \boldsymbol{W_k}$, is generally adopted. It is also assumed that some nodes of the graph may be both input nodes and out nodes (i.e. $\boldsymbol{Z} \cap \boldsymbol{W_k} \neq \emptyset$). For the case discussed in the paper it was assumed that the service is executed when from node $\boldsymbol{z_1}$ vehicles reach the predetermined nodes from $\boldsymbol{w_{k1}}$ to $\boldsymbol{w_{k4}}$, and, similarly, from $\boldsymbol{z_2}$ to corresponding from $\boldsymbol{w_{k2}}$ to $\boldsymbol{w_{k6}}$.

For the needs of research and to present the calculation possibilities a model of network connections structure in the form of an undirected graph is proposed, as shown in figure 2. For the analysis, estimated values of reliability indices of the connections of this structure, on condition of nodes availability, have been adopted.

The values of these reliability characteristics reflect the effects of all kinds of external influences on the system considered. External factors may be random or may be controlled by system operators. The estimation of the values of the characteristics is statistical and is the result of observation of the actual operating systems. Therefore, the article does not systematise the difficulties occurring within the transport network in the urban area, treating the obtained characteristics as a comprehensive representation of all external interactions.

Next, modifications in the structure were introduced by increasing the number of connections between the nodes, marked with dotted lines. The scheme of the structure with modifications is shown in figure 3.

For the modified structure, calculations were done again and the resulting changes in the network reliability were identified. In the analysis it was assumed that the network reliability measure, resulting from the factorisation algorithm, is the probability of the existence of a connection between a selected depot (Z1 or Z2) and all the end points of routes starting from the given depot. For depot Z1 the end points are WK1-WK4, and for depot Z2 the end points are WK2-WK6). In other words, it is the probability of executing all the scheduled trips from the given depot (reaching the destination point).

Figure 4 presents the network reliability measures discussed above, calculated for a standard network (Fig. 2) and modified one (Fig. 3), at various reliability values of connections between nodes: R=0,998; R=0,95; R=0,92 and R=0,72. To carry out calculations a

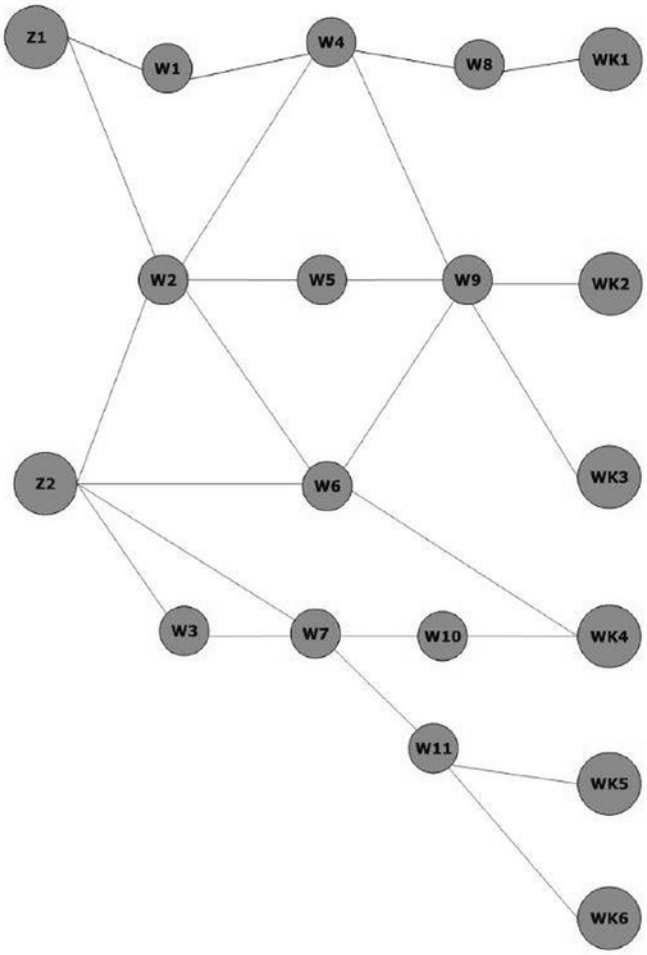


Fig. 2. Scheme of urban transport network structure

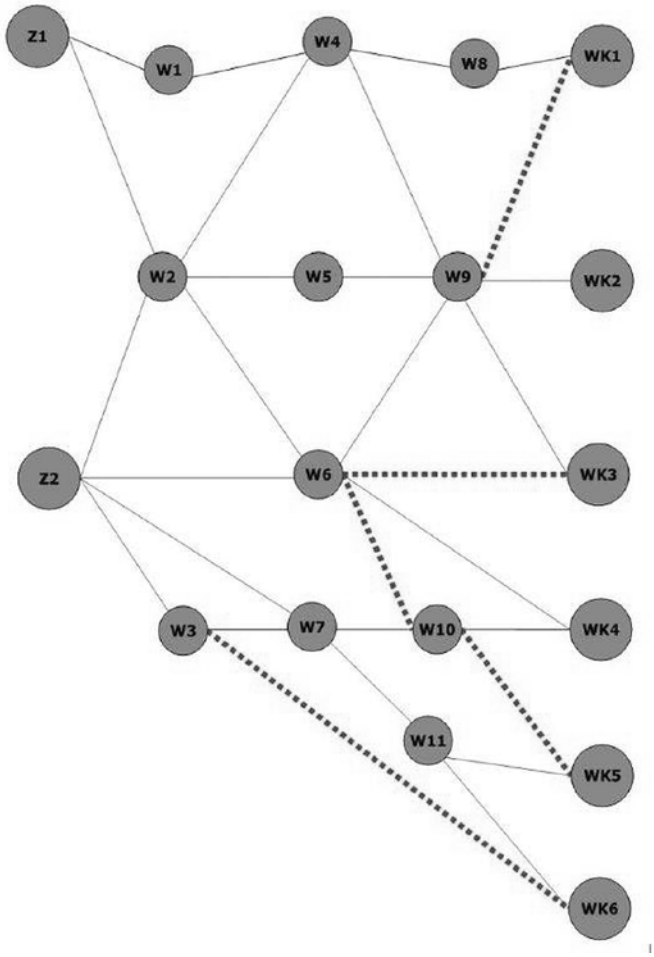


Fig. 3. Scheme of urban transport network structure modified

computer program was used, written in C language, which performs calculations in accordance with the presented procedure of the factorisation algorithm.

As can be noticed, only at high reliability of connections between nodes (R=0,998) the introduced modifications do not change significantly the network reliability (relative change for Z1 is 0,8% and for Z2 is also 0,8%). In the other cases, the lower the reliability of the connections the larger the change of network reliability resulting from modification. In the extreme case (R=0,72) the relative changes are, respectively, for Z1 65,2% and for Z2 69,8%.

Additionally, an analysis of the reliability of a network with failed nodes was performed. Node failure is treated as a failure of all the connections associated with the given node.

Consequently, the so-called critical nodes can be distinguished, i.e. nodes whose failure results in the lack of possibility of performing at least one of the scheduled trips from the given depot to end points of routes. In the case of a standard network (Fig. 2) these nodes are W4, W8, W9, W6, W7 and W11. In the case of a modified network (Fig. 3) the number of critical nodes is significantly reduced, and they are only W8 and W6.

Taking into account the failures of selected nodes (W4, W2, W10), which apart from W4 in the analysed network are not critical nodes, for a standard network the results obtained are shown as histograms (Figs 5 and 6).

The limit values in the form of relative percentage are:

– for Z1 failure of node W4 reduces the reliability by 100% in all cases (it is a critical node),

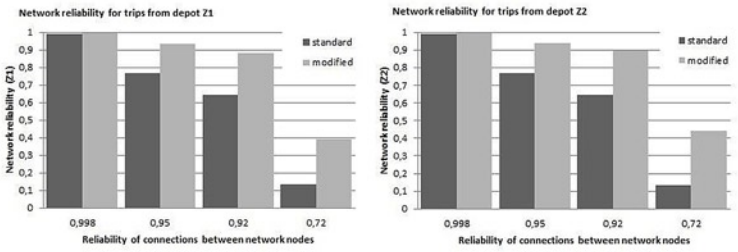


Fig. 4. Network reliability for trips from depots Z1 and Z2

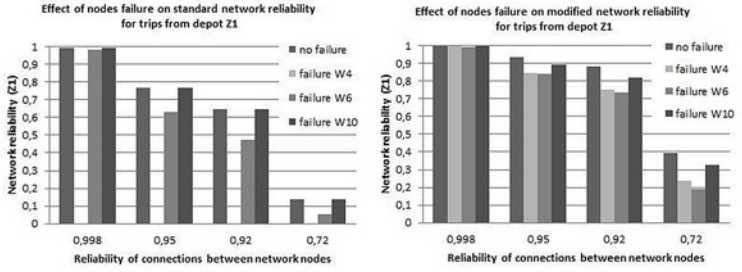


Fig. 5. Effect of nodes damage on network reliability for trips from depot Z1

– for Z2 failure of node W4 reduces the reliability in extreme cases by ≈0% for R=0,998 and by 7,7% for R=0,72,
– for Z1 failure of node W2 reduces the reliability in extreme cases by 0,8% for R=0,998 and by 61,9% for R=0,72,

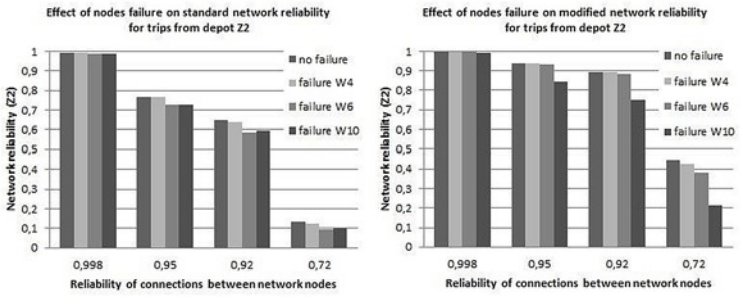


Fig. 6. Effect of nodes damage on network reliability for trips from depot Z2

– for Z2 failure of node W2 reduces the reliability in extreme cases by 0,2% for R=0,998 and by 30,7% for R=0,72,
– for Z1 failure of node W10 does not reduce reliability,
– for Z2 failure of node W10 reduces the reliability in extreme cases by 0,2% for R=0,998 and by 25,5% for R=0,72.

In an analogous way, the results for a modified network are (Figs 4 and 5):

– for Z1 failure of node W4 reduces the reliability in extreme cases by 0,4% for R=0,998 and by 38,9% for R=0,72,
– for Z2 failure of node W4 reduces the reliability in extreme cases by ≈0% for R=0,998 and by 3,7% for R=0,72,
– for Z1 failure of node W2 reduces the reliability in extreme cases by 0,4% for R=0,998 and by 51,6% for R=0,72,
– for Z2 failure of node W2 reduces the reliability in extreme cases by ≈0% for R=0,998 and by 14,5% for R=0,72,
– for Z1 failure of node W10 reduces the reliability in extreme cases by 0,2% for R=0,998 and by 16,8% for R=0,72,
– for Z2 failure of node W10 reduces the reliability in extreme cases by 0,4% for R=0,998 and by 51,5% for R=0,72.

5. Conclusions

The assessment of network systems reliability requires the use of specialised calculation methods which are not covered by the classic theory of reliability.

These methods include the factorisation algorithm which is useful in fast assessment of changes of reserving effectiveness according to the number and location of power supply points, in the selection of an optimum number of power sources and their location so as to ensure the high reliability of power supply of the analysed part of the networks, in the design of network taking into account the reliability and cost effectiveness of reserving. Its application enables the monitoring of network reliability changes when the structure is modified (by e.g. extending the connections), which will be practically useful for the designers of new networks as well as in modernisation and development projects of the existing networks. The factorisation algorithm used in the above analysis enables specification of reliability of networks of any configuration.

Instead of the classic reliability indices and characteristics the application of failure stream parameter in reliability assessment, which not only does not reduce the possibility of modernisation of the examined systems, on the contrary, it extends the utilitarian character of reliability analyses performed.

Multiple and fast repetitions of calculations for different variants of network modification require the algorithm to be saved as a computer program.

Network reliability assessment extended by an analysis of economy and operation related risk analysis can be useful for transport network designers, in modernisation and development projects of the existing networks.

The article presents reliability evaluation of the selected parts of a transportation network, assuming that the possibility of completing journeys between selected network nodes is an important criterion for users (passengers) of public transport. The possibilities of occurrence of failures between connections in network nodes are observed in real operating conditions and constitute a serious impediment to the functioning of public transport. For this reason, the presented reliability analysis should be useful for urban transport companies in the planning and development of transportation systems.

The factorisation method is presented in the literature presented below and it was not characterised in a detailed way, presenting only the results of calculations because it would cause a significant increase in the volume of the article, which aims to present the methodology for estimating the reliability of network systems, and not the factorization method.

References

1. Borowiecki R, Kaczmarek J, Magiera J, Młynarski S. Eksploatacja taboru szynowego komunikacji miejskiej. Niezawodność, jakość, ekonomika. Kraków: Wydawnictwo Akademii Ekonomicznej w Krakowie, 2004.
2. Brons M, Nijkamp P, Pels E, Rietveld R. Efficiency of urban public transit: A meta analysis. Transportation 2005; 32: 1-21, https://doi.org/10.1007/s11116-004-0939-4.
3. Bryniarska Z. Ocena satysfakcji pasażerów z innowacyjnego rozwiązania w miejskim transporcie zbiorowym w Krakowie. Transport miejski i regionalny 2017; 11: 5-9.
4. Carlier J, Lucet C. A decomposition algorithm for network reliability evaluation. Discrete Applied Mathematics 1996; 65: 141-156, https://doi.org/10.1016/0166-218X(95)00032-M.
5. Cats O, Koppenol GJ, Warnier M. Robustness assessment of link capacity reduction for complex networks: Application for public transport systems. Reliability Engineering and System Safety 2017; 167, https://doi.org/10.1016/j.ress.2017.07.009
6. Crainic TG, Ricciardi N, Storchi G. Models for Evaluating and Planning City Logistics Systems. Transportation Science 2009; 43(4): 432-454, https://doi.org/10.1287/trsc.1090.0279.
7. Heinrich M, Lenkiewicz W. Erhöhung und Auswertung des Zuverlässigkeitsniveaus von tribologischenPaarungenimZeitraum der AnfangsnutzungeineseinzelnenObjekts. Tribologie und Schmierungstechnik 2002, 3.
8. Heinrich M. Badania eksploatacyjne jednostkowych złożonych obiektów technicznych w celu podniesienia ich niezawodności konstrukcyjnej. Zagadnienia Eksploatacji Maszyn 1993; 1-2 (93-94): 125-133.
9. Holmgren J. The efficiency of public transport operations - An evaluation using stochastic frontier analysis. Research in Transportation Economics 2013; 39 (1): 50-57, https://doi.org/10.1016/j.retrec.2012.05.023.
10. Jan R-H. Design of reliable networks. Computers & Operations Research 1993; 20: 25-34, https://doi.org/10.1016/0305-0548(93)90093-X.
11. Jong GDe, Cheung F. Stochastic frontier models for public transport. Meersman H, van de Voorde E, Winkelmans W eds. World Transport

Research: Selected Proceedings of the 8th World Conference on Transport Research. Transport Modes and Systems. New York: Pergamon, 1999; 1: 373-386.
12. Karpiński J, Korczak E. Metody oceny niezawodności dwustanowych systemów technicznych. Omnitech Press, 1990.
13. Koide T, Shinmori S, Ishii H. Topological optimization with a network reliability constraint. Discrete Applied Mathematics 2001; 115: 135-149, https://doi.org/10.1016/S0166-218X(01)00221-9.
14. Levaggi R. Parametric and non-parametric approach to efficiency: The case of urban transport in Italy. StudiEconomici 1994; 49: 67-88.
15. Lucet C, Manouvrier J-F. Exact methods to compute network reliability, in Statistical and Probabilistic Models in Reliability. Ionescu DC and Limnios N eds. Boston: Birkhauser, 1999: 279-294, https://doi.org/10.1007/978-1-4612-1782-4_20.
16. Malasek J. Metoda oceny dostępności i atrakcyjności przystanków miejskiego transportu zbiorowego. Transport miejski i regionalny 2017; 9: 26-32.
17. Młynarski S. Niezawodność strukturalna w systemach logistycznych ratownictwa. Logistyka [CD] 2014; 4: 871-880.
18. Nolan JF. Determinants of productive efficiency in urban transit. Logistics and Transportation Review 1996; 32: 319-342.
19. Oprzędkiewicz J, Młynarski S. Problems of economic safety and intelligence system and reliability of global systems. In Enterprises in the face of 21st century challenges. Development - management - entrepreneurship. Borowiecki R & Jaki A eds. Cracow: Cracow University of Economics, 2008: 307-315.
20. Oprzędkiewicz J. Komputerowa metoda oceny niezawodności systemów technicznych. Lublin: LTN, 1997.
21. Page LB, Perry JE. A practical implementation of the factoring theorem for network reliability. IEEE Transactions on Reliability 1988; 37: 259-267, https://doi.org/10.1109/24.3752.
22. Paszkowski J, Kucharski R. Paradoksy przepustowości miejskiej sieci drogowej i sposoby ich odwzorowania w modelu czterostadiowym. Transport miejski i regionalny 2017; 10: 5-11.
23. Pilch R, Szybka J. Koncepcja zastosowania algorytmu faktoryzacji do oceny niezawodności ciągów komunikacyjnych - Application of factoring algorithms for estimation of road network reliability. Problemy Eksploatacji - Maintenance Problems 2007; 2: 129-136.
24. Pilch R. Factorisation algorithm-based metod used for the calculation of network system's reliability - Metodyka wyznaczania niezawodności układów sieciowych w oparciu o algorytm faktoryzacji. Zagadnienia Eksploatacji Maszyn - Scientific Problems of Machines Operation and Maintenance 2011; 4 (168): 45-57.
25. Pilch R. Reliability evaluation of networks with imperfect and repairable links and nodes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (1): 19-25, https://doi.org/10.17531/ein.2017.1.3.
26. Ramirez-Marquez JE, Rocco CM. All-terminal network reliability optimization via probabilistic solution discovery. Reliability Engineering and System Safety 2008; 93: 1689-1697, https://doi.org/10.1016/j.ress.2008.01.001.
27. Rymarz J, Niewczas A, Krzyżak A. Comparison of operational availability of public city buses by analysis of variance. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (3): 373-378, https://doi.org/10.17531/ein.2016.3.8.
28. Smalko Z. Modelowanie eksploatacyjnych systemów transportowych. Radom: ITE, 1996.
29. Smolnik M. A conception of modernising LEMACH designing methods using TRIZ instruments. Souchkov V, Kässi T. eds. TRIZfest-2014. Theories and applications. The 10th international conference: September 4-6, 2014, Prague, Czech Republic. Conference proceedings. noxville: International TRIZ Association - MATRIZ, 2014.
30. Smolnik M. Koncepcja systemu informacyjnego wspomagającego kierowanie eksploatacją złożonych odnawialnych obiektów technicznych. Logistyka 2015; 5: 1307-1312.
31. Wiśniewski P. Projekt i implementacja systemu wspomagania decyzji dla zarządzania kryzysowego transportem miejskim. Transport miejski i regionalny 2017; 11: 10-16.
32. Wood RK. Factoring algorithms for computing K-terminal network reliability. IEEE Transactions on Reliability 1986; 35 (3): 269-278, https://doi.org/10.1109/TR.1986.4335431.
33. Zamojski W. Niezawodność i eksploatacja systemów. Wrocław: Politechnika Wrocławska, 1981.

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

Robert PILCH
Maksymilian SMOLNIK
Jan SZYBKA
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,

Article citation info:
PIEGDOŃ I, TCHÓRZEWSKA-CIEŚLAK B, EID M. Managing the risk of failure of the water supply network using the mass service system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 284–291, http://dx.doi.org/10.17531/ein.2018.2.15.

Izabela PIEGDOŃ
Barbara TCHÓRZEWSKA-CIEŚLAK
Mohamed EID

MANAGING THE RISK OF FAILURE OF THE WATER SUPPLY NETWORK USING THE MASS SERVICE SYSTEM

ZARZĄDZANIE RYZYKIEM AWARII SIECI WODOCIĄGOWEJ Z WYKORZYSTANIEM SYSTEMU MASOWEJ OBSŁUGI

The aim of this paper is to analyse the functioning of the repair brigades in the process of failure removal in the water distribution subsystem using the mass service system (MSS). An example is presented using queuing model which takes into account notifications with various scheduling algorithms of failures occurring to the system. The functioning analysis of mass service system can be used in the optimization of the repair teams' actions and in the management of water supply companies.

***Keywords**: water distribution subsystem, queuing theory, management, modelling.*

Celem pracy jest analiza funkcjonowania brygad naprawczych w procesie usuwania awarii w podsystemie dystrybucji wody przy użyciu systemu masowej obsługi (SMO). Przykład został przedstawiony przy użyciu modelu kolejek, który uwzględnia zgłoszenia napływające do systemu z różnymi algorytmami planowania awarii. Funkcjonująca analiza systemu masowego świadczenia usług może być wykorzystana w optymalizacji działań zespołów naprawczych oraz w zarządzaniu przedsiębiorstwem wodociągowym.

***Słowa kluczowe**: podsystem dystrybucji wody, teoria kolejek, zarządzanie, modelowanie.*

1. Introduction

The theory of mass service called interchangeably the queuing theory [15, 16] is widely used in the search for mathematical models (analytical) allowing the most precise description of the services related to all branches of industry. The queuing theory is used in the analysis and description of the phenomena in which there is a problem of mass customer service [7, 14, 17, 32]. The main problem with the practical use of the theory of mass service system (MSS) is to determine the optimal decisions at random arriving queries (notifications, events). A pioneer in this field was the Danish mathematician A.K. Erlang. Erlang published studies on the load of call centres in 1909 [9]. In 1917, he presented the formulas for the probability of call blocking, called the Erlang model [8]. According to [8, 9, 27] the MSS can be classified in two categories, as follows:

– systems with exponential service time and many servers,
– systems constant service time and a single server.

Other equally important works are associated with the name David G. Kendall. In the years 1951 and 1953 Kendall published the works on the queuing systems in which he systematized the mass service systems using the so-called Kendall's notation. Thanks to those works Kendall is considered as the founder of the science of mass service [15, 16]. Since then many valuable works, presenting the theoretical basis of various queuing models, were published [6, 10-12, 14, 17-19, 21, 22, 32, 40]. Nowadays, the queuing models and the theory of queues are constantly applied and implemented in various technical systems [1, 2, 5, 7, 13, 25, 26, 39]. It is also possible to link MSS with the modelling of adverse events occurring in systems included in critical infrastructures [3, 4, 20, 23, 24, 28-31, 33-38, 41, 42].

The main task of the theory of queues is to optimise the waiting time before satisfying the arriving queries. Solving the tasks for the MSS requires knowledge of two basic characteristics: queries (notifications) arrival rate and the waiting time before supplying the service.

In the water distribution subsystems (WDS) the MSS is a trade-off platform for random competition between "needs of service" and "servers". In the WDS, the needs of service can be represented by the notification about failures occurrence, while as the servers can be represented by the repair brigades which are capable of handling these notifications. The ultimate target is to supply water with the required pressure, adequate quality and quantity, to all recipients. The exploitation of the water supply system mainly includes:

- supervision, i.e. activities aim at getting information about the state of the system and its current changes through inspections,
- genesis, i.e. analysing the causes that led to the occurrence of a particular state,
- diagnosis, i.e. inference about the state of the system's components on the basis of the results of performed examination,
- forecasting, i.e. predicting the conditions of the system or its components in the future,
- water pressure measurements,
- sampling for microbiological, physical and chemical examinations,
- flushing water mains,
- repairs in the water supply network,
- patrolling the exploitation area.

Exploitation of water supply network requires not only maintaining its operation and its proper management, but also restoring its technical capability and utility.

The aim of the study is to analyse the functioning of the repair brigades in the process of failure removal in the water distribution system using the mass service models. The analysis of the MSS functioning can be used, amongst others, in the management of water supply companies. Moreover, the article presents the topic of using the mass service system in a water supply system. It is noteworthy that there are few publications in the field. However, a wide application can be seen especially for telecommunications systems. The novelty

of presented analysis is the fact, that notifications with priorities for failures arriving to the WDS have been given. The presented analysis can significantly contribute to increasing the efficiency of the repair brigades and thus to increasing the reliability of water supply to all water recipients.

2. Research methodology – The Queuing Systems

2.1. Classification of queuing systems and basics models

We can talk about MSS when, on one hand, we have the notifications of events which arrive at the system with specific intensity and which require to be handled, and, on the other hand, there are servers capable of handling these notifications. Sometimes, however, not all notifications will be handled so there is a possibility to cancel the notification.

Queuing systems can be classified [27]:

- the algorithm (the method) for the inflow of notifications (input stream characteristics). Input stream can be characterized by the average number of notifications per time unit, or by the mean time between the arrivals of two successive notifications. The algorithm can be deterministic (notifications arrive in regular intervals) or stochastic (the mean time between notifications is the expected value of a random variable - the time between successive notifications),
- the way of handling notifications by servers. The server can be busy when the notification enters the system, then we assume the mean time needed for handling a single notification or there is a lack of notifications at server. The service algorithm can be deterministic (time to handle the notification is constant) or stochastic (we can distinguish different time distribution for handling the notification),
- the task scheduling algorithm (rules for selecting notifications from the queue to be handled by server) which can be summarised as:
  - FIFO (First in, first out) - first notification is handled first,
  - LIFO (Last in, first out) - last notification is handled first,
  - SIRO (Service in random order) - "random" selection, regardless of the arrival order,
  - priority scheduling – notifications with higher priority are selected to be handled as the first, regardless how many notifications with lower priority are in the system:
  - the absolute priority,
  - the relative priority,
- the number of places in the queue,
- the number of service channels.

The MSS operations can be modelled as stochastic processes [8, 10, 12, 15, 16, 41, 42]. To mark different MSS types the Kendall's notation is commonly used and the system can be described by some parameters: $A/B/\mathrm{r}:(L,N)$ where:

$A = T_p$ – the distribution of the random variable T_1, i.e. the time between successive notifications,
$B = T_n$ – the distribution of the random variable T_2, i.e. the distribution of service time,
r – the number of servers,
L – the number of places in the queue,
N – the size of serviced population.

If L and N in the notation are omitted, it means that there are infinitely large (∞,∞).

It was assumed that the probability distributions of the time intervals between notifications of needs of service and in service are exponential distributions. Furthermore, the functions T_p and T_n are independent. The number of places in the queue and the number of serviced people are infinite.

The parameters of the process are [41]:

λ – parameter of exponential distribution of the random variable T_p, which is the intensity of the inflow of notifications to the MSS:

$$\lambda = \frac{1}{T_p} \tag{1}$$

μ – parameter of exponential distribution of the random variable T_n, repair rate :

$$\mu = \frac{1}{T_n} \tag{2}$$

k – the number of notifications arriving to the MSS (number of failures),
r – the number of servers (number of repair brigades), $r \geq 1$,
ρ – utilization rate of the MSS [15]

$$\rho = \frac{\lambda}{\mu} \tag{3}$$

To avoid the so-called jamming the condition of the inequality must be fulfilled:

$$r \geq \rho = \frac{\lambda}{\mu} \tag{4}$$

The probability of state P_k (t) in which the WDS at time interval t has k notifications, the intensity of notifications $\lambda_k(t)$ and the intensity of service $\mu_k(t)$ are calculated per unit of time and are dependent on the number of notifications at time interval t (or interval of time). With the stationary nature of the process and other assumptions we obtain: $P_k(t) = P_k$, $\lambda_k(t) = \lambda_k$, $\mu_k(t) = \mu_k$. The system of equations describing the process for a stationary process is as follows [41, 42]:

$$\begin{cases} \lambda_z P_0 = \mu_z P_1 \\ (\lambda_z + k\mu_z) P_k = \lambda_z P_{k-1} + (k-1)\mu_z P_{k+1} & k > r \\ (\lambda_z + r\mu_z)\, P_k = \mu_z P_{k-1} + r\mu_z P_{k+1} & k \geq r \end{cases} \tag{5}$$

The distribution of the probabilities of the number of damaged elements given by A.K. Erlang is:

$$P_k = \begin{cases} \dfrac{M!}{k!\,(M-k)!} \cdot \rho^k \cdot P_0\,; & k = 0,1,...,r \\ \dfrac{M!}{r!\, r^{k-r} \cdot (M-k)!} \rho^k \cdot P_0; & k = r, r+1, ...M \end{cases} \tag{6}$$

where :

$$P_0 = \frac{1}{[1 + \sum_{k=1}^{M} (\rho_1 \cdot \rho_2 \cdot ... \rho_k)]} \tag{7}$$

wherein:

$$\rho_k = \frac{\lambda_{k-1}}{\mu_k}; \quad k = 1,2,...,M \tag{8}$$

The probability that during the stationary period the system is in the k-th state takes the form:

$$P_k = \begin{cases} \frac{\rho^k}{k!} \cdot P_{01} & k < r \\ \frac{\rho^k}{r! \cdot r^{k-r}} \cdot P_{01} & k \geq r \end{cases} \tag{9}$$

The intensity of transitions in the k-th state is as follows:

$$\mu_k = \begin{cases} k \cdot \mu_z; & k = 0,1,2,...,r \\ r \cdot \mu_z; & k = r, r+1,...,M \end{cases} \tag{10}$$

$$\lambda_k = (M-k) \cdot \lambda_z; \quad k = 0,1,...,M \tag{11}$$

where:

μ_k – the intensity of service (service rate) in the k-th state,
λ_k – the intensity of notification inflow (arrival rate) in the k-th state.

The average number of notifications in the MSS is:

$$E(N_w) = \sum_{k=1}^{r+L_w} kP_k \tag{12}$$

The average number of notifications in the queue is:

$$E(U_w) = E(U_w) = \sum_{k=1}^{L_w} kP_{r+k} \tag{13}$$

The average number of free repair brigades is:

$$E(O_r) = \sum_{k=0}^{r-1} (r-k) \cdot P_k \tag{14}$$

There is the equality:

$$E(N_w) - E(U_w) + E(0_r) = r \tag{15}$$

The average waiting time for service is:

$$\mathrm{E}(\mathrm{T_w}) = E(T_w) = \frac{E(U)}{\lambda_z} \tag{16}$$

The average time when notification is in the MSS is:

$$E(T_z) = E(T_w) + \frac{1}{\mu_z} \tag{17}$$

The average waiting time for service when all the repair brigades are busy is [41]:

$$E(T_v) = \frac{E(T_w)}{P_{k \geq r}} \tag{18}$$

The number of repair brigades for the WDS should be:

$$r = r_{min} + r_{dod} \tag{19}$$

where:

r_{min} – the number of repair brigades necessary to avoid a queue or blocking the queue, which is defined by the condition (4),
r_{dod} – the number of additional repair brigades, which may be adopted from the condition for the required system reliability K_w [41].

2.2. Models of service with priorities

Reliability and safety of the WDS operation can be also considered assuming the priority MSS model. Among the priorities one can distinguish the absolute priority and the relative priority.
The notification arriving to the system gets the absolute priority if handling the notification causes that handling any other notification would be interrupted. However, the notification has the relative priority if it does not interrupt the handling of the other notifications. Giving priority to the notifications should be performed individually for each WDS taking into account its specificity.

The intensity of the notifications (arrival rate) of the first kind is λ_{z1} and the notifications of the second kind is λ_{z2}. It is assumed that the service time for both types of notifications is the same and that the intensity of service (service rate) is then equal to μ_z. Therefore it can be written [39]:

$$\lambda_z = \lambda_{z1} + \lambda_{z2} \tag{20}$$

$$\rho_i = \frac{\lambda_{zi}}{\mu_z}; \quad i = 1,2 \tag{21}$$

It can consider the cumulative probability $P\{N_1(t) = k_1, N_2(t) = k_2\}$, $N_1(t)$ and $N_2(t)$ - the number of notifications of the first and second kind arriving to the MSS at time interval t. It can be assumed that streams of the first and second kind failures are stationary streams (Poisson). In such case the probability that in a short period of time there will be more than one failure has small value of higher order than Δt [14].

When the first notification type, has the absolute priority, the second type notification has to wait and it is not possible to serve the second kind notification as the first one. Handling the notifications with the absolute priority takes place independently of the notifications with the relative priority. To describe handling the first kind notifications in the WDS the following model can be used: $A/B/\mathrm{r}:(L,N)$. The expected value of notifications that are in the system [39]:

$$E(N_{w1}) = \frac{\rho_1}{1-\rho_1} \tag{22}$$

The expected value of notifications waiting for service:

$$E(U_{w1}) = \frac{\rho_1^2}{1-\rho_1} \tag{23}$$

The probability that the server is free:

$$P_1 = 1 - \rho_1 \tag{24}$$

The expected value of the waiting time for service:

$$E(T_{w1}) = \frac{\lambda_1}{\mu \cdot (\mu - \lambda_1)} \tag{25}$$

The expected value and the waiting time for service if server is busy:

$$E(T_{V1}) = \frac{1}{(\mu_z - \lambda_{z1})} \tag{26}$$

The number of the second kind notifications can be calculated according to the formula [22]:

$$E(N_{w2}) = \frac{\rho_2}{1 - \rho_1 - \rho_2}[1 + \frac{\rho_1}{1 - \rho_1}] \tag{27}$$

The order of handling the notifications of one kind can be according to the following rules:

- first arrived - first served,
- last arrived - first served,
- random selection of service.

Handling the notifications of the first and second kind can be performed independently if:

$$\rho_1 + \rho_2 \leq r \tag{28}$$

The Erlang formula for both streams of notifications takes the form [10]:

$$p_{k1,k2} = \frac{(\rho_1 + \rho_2)^{k_1 + k_2}}{(k_1 + k_2)! \sum_{i=1}^{M} (\rho_1 + \rho_2) \cdot \frac{i}{i!}} \tag{29}$$

$$\rho_i = \frac{\lambda_{zi}}{\mu_z} \tag{30}$$

where:

$k_1 + k_2 \leq M$, M – number of population in service, i - the number of different streams of notifications

$i = 1,2,$

$\lambda_{z \leq i} = \lambda_{z1} + ... + \lambda_{zi}$

Downtime brigades indicator can be estimate using the formula [39]:

$$z = \frac{E(O)}{r} \tag{31}$$

3. Application case

3.1. Research object

The daily production capacity at the end of 2011 amounted to 84000 m^3/d. The average daily production of treated water in Water Treatment Plans amounts to 37700 m^3/d. The system of collective water supply at the end of 2011 covered 184152 inhabitants. The main network is made of cast iron and steel pipes. The distribution network is constructed from cast iron, steel, PE and PVC pipes. The skeleton of water supply system consists of four mains transporting treated water from the second stage pumping station. 80% of the network is made in a closed system. In addition to the operation of the water supply system and water supply connections in the city operate also: emergency deep-seated intake with a capacity of 240 m^3/d, 32 water pumping stations, 12 clean water compensating tanks with a total capacity of 34100 m^3/d, 187 public wells.

Calculation example for the water distribution subsystem was based on the list of failures in the water supply system and water supply connections. The data taken for the analysis are summarized in Table 1.

Figure 1 presents percentage distribution of the number of failures in the water supply network and water connections.

In the city, water and sewage emergency services operate six repair brigades. At the beginning of a shift, a task is assigned to each brigade based on the notification from the water and sewer emergency services. The order of handling failures is set every time (handling

Table 1. Failures of water network and water connections over the years 2005-2012

Years	Type of pipes					
	Main		Distribution		Water connections	
	Length (km)	Number of failures	Length (km)	Number of failures	Length (km)	Number of failures
2005	49,5	54	350,5	108	283,8	83
2006	49,5	45	384,4	136	287,7	117
2007	49,5	51	443,5	114	315,8	90
2008	49,5	29	447,7	106	322,8	83
2009	49,8	38	468	114	323,2	65
2010	49,8	39	490,5	114	323,8	102
2011	49,8	52	504,1	113	323,8	134
2012	49,8	55	520,5	109	323,8	119

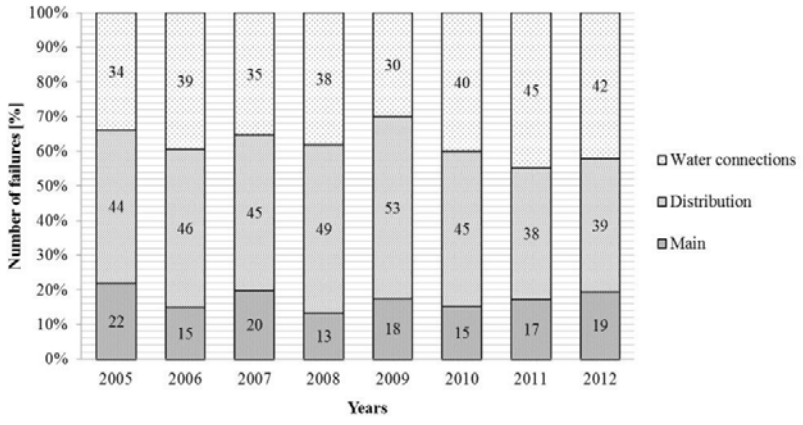


Fig. 1. The number of failures in the water supply network and water connections, in %

the failures is determined based on their priority service given by the master of the brigade). The brigades with the materials necessary to handle the failure are carried to the designated place. The application example were analysed for the brigades: 4 (service of water network) + 2 (service of water connections).

3.2. Results of modelling

a) Service of water network. System works without the priority of notifications (all notifications have the same priority).

The MSS model type A/B/r:(∞,∞) was adopted. Assumptions for modelling based on exploitation data:

- the mean time between failures $Tp_{śr}$ (d), $Tp_{śr}$ = 2,23 d
- unit intensity of notifications (arrival rate) λ_z = 0,448 1/d
- the mean time of repair $Tn_{śr}$ (d), $Tn_{śr}$ = 0,132 d
- repair rate μ_z = 7,57 1/d
- mass service system utilization rate ρ = 0,059
- the number of repair brigades, r = 4
- minimum number of repair brigades r_{min} = 1

The results of the calculations are presented in Table 2.

b) Service of water network. System works without the priority of notification (all notifications have the same priority).

Assumptions for the modelling were the same as in section 1a with the change in the number of repair brigades according to the condition r ≥ ρ. To avoid a queue or queue blocking. The number of necessary repair brigades is r_{min} = 1, where r_{min} is the smallest natural number satisfying the above inequality. The results of the calculations are presented in Table 3.

c) Service of water network. System works with the absolute priority of notification.

For the calculations it was assumed, that the notifications with the absolute priority will arrive to the system.

c1) Assumptions for modelling of the first kind of notifications (based on exploitation data):

- the number of notifications with priority k_1 is assumed to be 95% of all the notifications arriving to the MSS.
- k_1 = 156,
- the mean time between failures $T_{pśr}$ (d), $T_{pśr}$ = 2,34 d
- unit intensity of notifications inflow (arrival rate) λ = 0,427 1/d
- the mean time of repair $T_{nśr}$ (d), $T_{nśr}$ = 0,132 d
- repair rate μ_z = 7,57 1/d
- mass service system utilization rate ρ_1 = 0,056

Table 2. Summary of results

The number of notifications k	Number of free repair teams r_k	The probability of state P_k	E(N)	E(U)	E(O)	z
	SMO1a, r = 4, ρ = 0,059					
0	4	0,8416	0,1676	0,0002	3,8326	0,958
1	3	0,1494				
2	2	0,0088				
3	1	0,0002				

Table 3. Summary of results

The number of notifications k	Number of free repair teams r_k	The probability of state P_k	E(N)	E(U)	E(O)	z
	SMO1b, r = 1, ρ = 0,059					
0	1	0,8335	0,1861	0,0010	0,8149	0,814
1	0	0,1480				
2	0	0,0175				
3	0	0,0010				

Table 4. Summary of results for the first type of notifications

The number of notifications k	Number of free repair teams r_k	The probability of state P_k	E(N)	E(U)	E(O)	z
	SMO1c, r = 4, ρ = 0,056					
0	4	0,8482	0,1603	0,0002	3,8399	0,960
1	3	0,1435				
2	2	0,0081				
3	1	0,0002				

Table 5. Summary of results for the second type of notifications

The number of notifications k	Number of free repair teams r_k	The probability of state P_k	E(N)	E(U)	E(O)	z
	SMO1c, r = 4, ρ = 0,003					
0	4	0,9913	0,0087	0,0000	3,9913	0,998
1	3	0,0086				
2	2	0,0001				
3	1	0,0000				

Table 6. Summary of the results for handling water supply connections.

The number of notifications k	Number of free repair teams r_k	The probability of state P_k	E(N)	E(U)	E(O)	z
	SMO2, r = 2, ρ = 0,032					
0	2	0,9108	0,0920	0,0000	1,9080	0,954
1	1	0,0864				
2	0	0,0028				
3	0	0,0000				

Table 7. Summary of the results for the whole work of repair brigades

The number of notifications k	Number of free repair teams r_k	The probability of state P_k	E(N)	E(U)	E(O)	z
	SMO3, r = 6, ρ = 0,116					
0	6	0,7191	0,3123	0,0011	5,6877	0,948
1	5	0,2507				
2	4	0,0291				
3	3	0,0011				
4	2	0,0000				
5	1	0,0000				
6	0	0,0000				

c2) Assumptions for modelling of the second kind of notifications (based on exploitation data):
- the number of notifications with priority k_2 is assumed to be 5% of all the notifications arriving to the MSS,
- $k_2 = 8$,
- the mean time between failures $Tp_{śr}$ (d), $Tp_{śr} = 45,63$ d
- unit intensity of notifications inflow (arrival rate) $\lambda_z = 0,022$ 1/d,
- the mean time of repair $Tn_{śr}$ (d), $Tn_{śr} = 0,132$ d,
- repair rate $\mu_z = 7,57$ 1/d,
- mass service system utilization rate $\rho_2 = 0,003$.

Since $\rho 1 + \rho 2 < r$ handling arriving notifications can be carried out independently.

In Table 4 and 5 the calculation results for the sequence of notifications of the first type and second type are presented.

d) Service of water connections

Assumptions for modelling based on exploitation data:
- the mean time between failures $Tp_{śr}$ (d), $T_{pśr} = 3,07$ d
- unit intensity of notifications inflow (arrival rate) $\lambda_z = 0,326$ 1/d
- the mean time of repair $Tn_{śr}$ (d), $Tn_{śr} = 0,097$ d
- repair rate $\mu_z = 10,31$ 1/d
- mass service system utilization rate $\rho = 0,032$
- the number of repair brigades r = 2

Table 6 summarizes the results of calculations for the work of brigades serving the water supply system connections.

e) Service of water network and water connections. Teamwork of all the brigades without priority.

Assumptions for modelling based on exploitation data:
- the mean time between failures $Tp_{śr}$ (d), $Tp_{śr} = 1,29$ d
- unit intensity of notifications inflow (arrival rate) $\lambda_z = 0,7751$/d
- the mean time of repair $Tn_{śr}$ (d), $Tn_{śr} = 0,150$ d
- repair rate $\mu_z = 6,67$ 1/d
- mass service system utilization rate $\rho = 0,116$
- the number of repair brigades r = 6

Table 7 summarizes the results of calculations for the work of brigades serving the water supply system and water connections.

On Figure 2 calculated average numbers of notifications in the MSS was presented.

After performing the calculations it is possible to check the reliability condition, in order to check whether the given MSS has the required level of reliability [41]:
- the availability index of one repair brigade: $K_g = 0,9923077$,
- the required level of MSS reliability: $K_w = 0,9965225$. The availability index of MSS takes the form:

$$K_g(SMO) = \sum_{i=r\min}^{r} \binom{r}{i} \cdot K_g{}^{i} \cdot \left(1 - K_g\right)^{r-1} \tag{31}$$

The reliability condition:

$$K_g(SMO) \geq K_w \tag{32}$$
$$K_g(SMO3) \geq K_w \rightarrow 0,9999999 \geq 0,9965225.$$

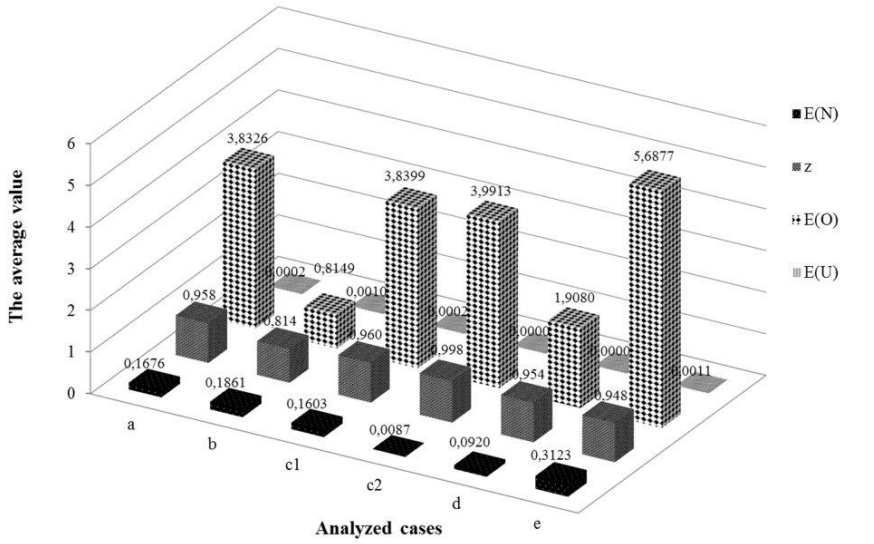


Fig.2. The comparison of calculated average numbers of notifications for MSS for water supply system

The required level of MSS reliability has been maintained because the condition (32) was met when the repair brigades worked together.

4. Conclusions

- The conducted calculations show that the most effective model of work of the water and sewage emergency service is the work of brigades in case e), when there is a teamwork of all the brigades without priority of failures. It results from obtaining the lowest outage index of repair brigades.
- The proposed methodology allows checking whether with the given number of repair brigades in the water supply company the jamming is not created while notifications are handled (idem).
- The fulfilment of the reliability condition with repair brigades working as a team causes that MSS has the required degree of reliability.
- The method can be used to search for solutions with the smallest possible number of repair brigades with maintaining the required level of reliability (minimum cost - maximum reliability).
- However, for the proper functioning of the emergency service the distribution of notifications arriving to the system should be predicted. The classification would predict the division of notifications on those that require immediate service and those whose handling is not so urgent. During such classification should be considered the economic aspects as well as the aspects related to safety of water supply to consumers.
- Analysing different variants of repair brigades work allows to make decisions concerning the improvement of work of water and sewer emergency service, which in consequence causes that the satisfaction of water consumers who use the service increases.
- The proposed method allows to analyse the repair brigades work in every water supply company, regardless of its size.
- The MSS analysis can be used in the management process in water supply companies in order to increase the efficiency in the process of management and making decisions related to operation of the water supply system.
- The article presents the topic of using the mass service system in a water supply system. It is noteworthy that there are few publications in the field. However, a wide application can be seen especially for telecommunications systems. The novelty of presented analysis is the fact, that notifications with priorities for failures arriving to the WDS have been given. The presented analysis can significantly contribute to increasing the efficiency of the repair brigades and thus to increasing the reliability of water supply to all water recipients.
- Service with priorities can be applied for different water recipients depending on water supply regime. For example, absolute priority can be attributed to failures on water pipes supplying in water health care institutions, nursing homes or nurseries. Water for firefighting purposes should have independent source of water. The presented model can be used in crisis management plans for urban agglomerations.

References

1. Artalejo J, Gomez-Corral A. Retrial queueing systems. Springer, Berlin 2008, https://doi.org/10.1007/978-3-540-78725-9.
2. Bolch G, Greiner S, De Meer H, Trivedi K. Queueing networks and Markov chains. Wiley & Sons, New York 2006, https://doi.org/10.1002/0471791571.
3. Boryczko K. Water age in the water supply network as health risk factor associated with collective water supply. Ecological Chemistry And Engineering A-Chemia I Inzynieria Ekologiczna 2016: 23 (1): 33-43, https://dx.doi.org/10.2428/ecea.2016.23(1)3.
4. Boryczko K, Piegdoń I, Eid, M. Collective water supply systems risk analysis model by means of RENO software. Safety, Reliability and Risk Analysis: Beyond the Horizon, Taylor & Francis Group, London 2014: 1987-1992.
5. Bose S. An introduction to queueing systems. Kluwer Academic Plenum Publishers, New York 2001.
6. Cinczyn A J. Roboty po matematiczeskoj teorii masowego obsługiwania. Fizmatgiz, Moskwa 1963.
7. Daigle J. Queueing theory with applications to packet telecommunication. Springer, New York 2005, https://doi.org/10.1007/b99875.
8. Erlang A K. Solution of some problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges. Elektrotkeknikeren 1917: 13: 5-13.
9. Erlang A K. The theory of probabilities and telephone conversations. Nyt Tidsskrift for Matematik B 1909; 20: 33-39.
10. Gniadenko B. W. Prioritietnyje sisiemy obsłużywanija. Uniwersytet Moskiewski, Moskwa 1973.
11. Gniadenko B W, Kowalenko I N. Introduction to queueing theory. Birkhauser, Boston 1991.
12. Gniadenko B W, Kowalenko I N. Wwiedienije w teoriu massowego obsługiwania. Nauka, Moskwa 1987.
13. Iwanejko R, Bajer J. Determination of the optimum number of repair units for water distribution systems. Archives of Civil Engineering 2009; 55 (1): 87-101.
14. Jaiswal N. Priority queues. Academic Press, New York 1969.
15. Kendall D G. Some problems in the theory of queues. Journal of the Royal Statistical Society B 1951; 13: 151-185.
16. Kendall D G. Stochastic processes occurring in the Theory of Queues and their analysis by the method of the imbedded Markov chain. The Annals of Mathematical Statistics 1953; 24 (3): 338-354.
17. Kleinrock L. Queueing systems (in Hungarian). Muszaki Kiado, Budapest 1975.
18. Klimow G P. Procesy masowej obsługa. WNT, Warszawa 1979.
19. Konig D, Stoyan D. Methoden der Bedienungstheorie. Akademie Verlag, Berlin 1976.
20. Kotowski A, Kaźmierczak B. Dimensioning of the sewerage system. Proceedings of EC Opole 2013; 7(1): 117-124.
21. Kozłow W W, Sołowiew A D. Optimalnoje obsłużiwanie wosstanawliwajemych system Izwiestia AN SSSR. Techniczeska Kibernetika 1978; (3-4).
22. Koźniewska I, Włodarczyk M. Modele odnowy, niezawodności i masowej obsługi, Państwowe Wydawnictwo Naukowe, Warszawa 1978.
23. Kutyłowska M. Modelling of failure rate of water-pipe networks. Periodica Polytechnica Civil Engineering 2015; 1: 37-43, https://doi.org/10.3311/PPci.7541.
24. Kutyłowska M. Neural network approach for failure rate prediction. Engineering Failure Analysis 2015; 47: 41-48, https://dx.doi.org/10.1016/j.engfailanal.2014.10.007.

25. Nelson R. Probability stochastic processes and queueing theory. The mathematics of computer performance modelling. Springer-Verlag, New York 1995, https://dx.doi.org/ 10.1007/978-1-4757-2426-4.
26. Nowacka A, Włodarczyk-Makuła M, Tchórzewska-Cieślak B, Rak J. The ability to remove the priority PAHs from water during coagulation process including risk assessment. Desalination and Water Treatment 2016; 57(3): 1297-1309, https://dx.doi.org/10.1080/19443994.2015.1030108.
27. Oniszczuk W. Metody modelowania. Wydawnictwo Politechniki Białostockiej, Białystok 1995.
28. Piegdoń I, Tchórzewska-Cieślak B. Methods of visualizing the risk of lack of water supply. Safety and Reliability: Methodology and Applications, Taylor & Francis Group, London 2014; 497-505.
29. Piegdoń I, Tchórzewska-Cieślak B, Szpak D. The use of geographical information system in the analysis of risk of failure of water supply network. Environmental Engineering V, Taylor & Francis Group, London 2017: 7-14.
30. Pietrucha-Urbanik K, Studziński A. Case study of failure simulation of pipelines conducted in chosen water supply system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19(3): 317-323, http://dx.doi.org/10.17531/ein.2017.3.1.
31. Rak J, Boryczko K. Assessment of water supply diversification using the Pielou index. Environmental Engineering V, London 2017: 53-58.
32. Solomon S. Simulation of waiting line systems Prentice Hall, Englewood-Cliffs, New York 1983.
33. Szpak D, Tchórzewska-Cieślak B. Water producers risk analysis connected with collective water supply system functioning, Dependability Engineering and Complex Systems. Advances in Intelligent Systems and Computing 470, Springer, Berlin 2016: 479-489, https://dx.doi.org/10.1007/978-3-319-39639-2_42.
34. Tchórzewska-Cieślak B, Rak J. Method of identification of operational states of water supply system. Environmental Engineering III, Taylor & Francis Group, London 2010: 521-526, https://doi.org/10.1201/b10566-83.
35. Tchórzewska-Cieślak B, Boryczko K, Eid M. Failure scenarios in water supply system by means of fault tree analysis. Advances in Safety, Reliability and Risk Management, Taylor & Francis Group, London 2012: 2492-2499.
36. Tchórzewska-Cieślak B, Boryczko K, Piegdoń I. Possibilistic risk analysis of failure in water supply network, Safety and Reliability: Methodology and Applications, Taylor & Francis Group, London 2014: 1473-1480.
37. Tchórzewska-Cieślak B, Pietrucha-Urbanik K, Urbanik M. Analysis of the gas network failure and failure prediction using the Monte Carlo simulation method. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(2): 254-259, https://dx.doi.org/ 10.17531/ein.2016.2.13.
38. Tchórzewska-Cieślak B, Szpak D. A proposal of a method for water supply safety analysis and assessment. Ochrona Środowiska 2015; 37(3): 43-47.
39. Tchórzewska-Cieślak B. Reliability of selected elements of the natural gas supply subsystem. PhD Dissertation, Cracow 2002.
40. White J A, Schmidt J W, Bennett G K. Analysis of Queueing Systems. Academic Press, New York 1975.
41. Wieczysty A, Iwanejko R, Lubowiecka T, Rak J. Określenie liczby brygad remontowych w podsystemie dystrybucji wody przy zastosowaniu modelu masowej obsługi. Gaz, Woda i Technika Sanitarna, 1990; 7.
42. Wieczysty A, Iwanejko R, Lubowiecka T, Rak J. Wyznaczenie liczby brygad remontowych obsługujących sieć wodociągową miasta. Zeszyty Naukowe Politechniki Rzeszowskiej. Budownictwo i Inżynieria Środowiska 1991; 12: 259-263.

Izabela PIEGDOŃ
The Faculty of Civil and Environmental Engineering and Achitecture
Rzeszow University of Technology
al. Powstańców Warszawy 12,
35-959 Rzeszow, Poland

Barbara TCHÓRZEWSKA-CIEŚLAK
The Faculty of Civil and Environmental Engineering and Achitecture
Rzeszow University of Technology
al. Powstańców Warszawy 12,
35-959 Rzeszow, Poland

Mohamed EID
CEA/DANS/DM2S/SERMA
CE Saclay, Bat. 470
F-91191 Gif sur Yvette Cedex, France

E-mails: piegi@prz.edu.pl, cbarbara@prz.edu.pl, mohamed.eid@cea.fr

Article citation info:
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.

Andrzej ŚWIDERSKI
Arkadiusz JÓŹWIAK
Roland JACHIMOWSKI

OPERATIONAL QUALITY MEASURES OF VEHICLES APPLIED FOR THE TRANSPORT SERVICES EVALUATION USING ARTIFICIAL NEURAL NETWORKS

EKSPLOATACYJNE MIARY JAKOŚCI POJAZDÓW W ZASTOSOWANIU DO OCENY USŁUG TRANSPORTOWYCH Z WYKORZYSTANIEM SZTUCZNYCH SIECI NEURONOWYCH*

Operational vehicle quality measures are an important element used to evaluate the performance of transport services. In practice, there are many methods involved in the operational evaluation of vehicles. They are characterized in this article. Artificial Intelligence methods, especially artificial neural networks, can also be successfully used for this purpose, and especially when deciding on quality assessment processes for machines, including motor vehicles. The use of methods to support decision-making based on facts is extremely important for the credibility and objectivity of the evaluation. These methods can also be used in relation to the use of vehicles in the assessment of transport services. The article presents the method of using artificial neural networks for the operational evaluation of vehicles used in freight transport services. The basis for the verification of the method was an experimental research carried out at a company making dairy products, cooperating with transport companies, supplying products for the production process. The results obtained from the operation of vehicles from the studied companies have confirmed, at the probability level of 99%, high efficiency of the proposed method in evaluating transport services using operational vehicle quality measures.

***Keywords**: vehicles operation, evaluation of transport services, quality measures, artificial neural networks.*

Eksploatacyjne miary jakości pojazdów są istotnym elementem wykorzystywanym do oceny realizacji usług transportowych. W praktyce mamy do czynienia z wieloma metodami związanymi z eksploatacyjną oceną pojazdów. Scharakteryzowano je w artykule. Metody sztucznej inteligencji, a zwłaszcza sztuczne sieci neuronowe, również mogą być z powodzeniem wykorzystane do tego celu, a zwłaszcza przy podejmowaniu decyzji w procesach oceny jakości maszyn, w tym pojazdów samochodowych. Zastosowanie metod, które pozwalają wspomagać proces decyzyjny na podstawie faktów jest niezmiernie istotne z punktu widzenia wiarygodności i obiektywności oceny. Metody te mogą być również wykorzystane w odniesieniu do eksploatacji pojazdów w zastosowaniu do oceny usług transportowych. W artykule przedstawiono metodę wykorzystania sztucznych sieci neuronowych do eksploatacyjnej oceny pojazdów wykorzystywanych w usługach transportowych towarów. Podstawę weryfikacji metody stanowiły badania eksperymentalne przeprowadzone w przedsiębiorstwie produkującym produkty mleczarskie, współpracującym z firmami transportowymi, dostarczającymi wyroby do produkcji. Uzyskane wyniki potwierdziły z 99-procentowym prawdopodobieństwem wysoką skuteczność proponowanej metody w dokonywaniu oceny usług transportowych z wykorzystaniem eksploatacyjnych miar jakości pojazdów.

***Słowa kluczowe**: eksploatacja pojazdów, ocena usług transportowych, miary jakości, sztuczne sieci neuronowe.*

1. Introduction

Operational quality measures of motor vehicles are used, among the others, to evaluate the performance of transport services. An important group of problems in making such an assessment is selection of the appropriate method. The operational evaluation of an object requires defining the measures (measurements, indicators) and the determining their values. The appropriate value allocation of the vehicles performance measures is one of the key criteria for the proper functioning of the whole transport system [9]. The numerical evaluation of the efficiency of the equipment is based on the values derived from the observation of the equipment during operation [10]. The variety of operational measures depends, of course, on the type of object (process), and usually these measures have different denominations and orders of scale, making them mutually incomparably [6, 11]. Comparing the measures describing an object (process) is only possible after normalization. Among the groups of technical objects' features relevant for their operational evaluation (determination of their measures and indicators) were distinguished, among the others [8]:

- technical condition of the object, being a measure of the ability to use the object over time,
- reliability in statistical terms,
- quality, understood as the ability of an object to meet specific needs,
- functionality describing the object in the sphere of human contact,
- efficiency characterizing the performance of an object,
- serviceability characterizing the object's suitability to be serviced,

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

- ability to be diagnosed characterizing the object's susceptibility to obtaining information on technical condition.

Determining the measures of the above mentioned groups of objects' features requires the use of mathematical models. The most commonly used models include reliability models [8, 14] overall equipment effectiveness models (OEE) and Key Performance Indicators models (KPI) [17].

The reliability model allows statistically determine operational measures. The reliability measure in this model is based on the reliability function defined as the probability of correct operation of the object in the assumed time [14]. In practice, the reliability models allow to define the indicators related to the operational objects in terms of technical and technical-organizational aspect.

The overall equipment effectiveness models focus the operation measures by the use of object's availability, efficiency, and quality of its operation. The Key Performance Indicators model (KPI) includes a collection of key performance and efficiency measures. These measures are specified in EN 15341: 2007 standard (Maintenance - Maintenance Key Performance Indicators). This standard contains 72 indicators, with a detailed interpretation of the components that they contain [15].

Determining the value of the measures of the objects' operational features allows to further evaluate the changes of these values at a certain time. Assessing an objects or processes involves making a decision. Decisions usually have to satisfy the whole set of needs of the decision maker, which makes it necessary to compare possible solutions, variants in terms of many criteria characterizing a given object or process. Hence making complex decisions requires the use of multi-criteria analysis methods (MCDM - Multi Criteria Decision Making). These methods play an important role, among the others, in the diagnosis of existing objects or organizational solutions [22]. Due to the fact that the features of objects or systems are usually expressed in different units of measure, their states can not be directly compared with each other. It is only the division of the set of features characterizing a given object (system), in terms of the desired tendencies of the formation of their values, enables the unification of the partial criteria and the comparing of these characteristics. Multi-criteria decision supporting methods can be divided into those stemming from the usefulness theories (UTA, UTASTAR, AHP, ANP, SMART) and methods based on overriding relationships (e.g. ELECTRE, PROMETHEE, ORESTE, REGIME), which indicate that due to a certain criterion one solution is „at least as good" as the other solution. The UTA (UTilités Additives) method is based on the principle of aggregation/division. It uses linear programming techniques to optimally define additive value/usefulness functions so that these functions are as consistent as possible with the preferences of the decision maker [24]. The development of UTA method is a UTASTAR method. It has additionally two error functions that denote the violation of the lower and upper ends of the usefulness function of the alternatives group by the k-th decision variant [24].

The Analytic Hierarchy Process (AHP) is a generic hierarchical approach to multi-criteria decision making that allows to combine quantified criteria with non-quantified ones and objectively measurable with subjective [11, 17, 18]. Modelling with a hierarchical analysis of the problem AHP is particularly useful where there is no known functional dependence between the elements of the decision making problem, described in form of the hierarchy of the factors, but the effect of the property data occurring and their practical effect, can be estimated. The extension of the AHP method is the ANP method (Analytic Network Process) method [1]. It can be applied to solve more complex decision problems. A system of components relevant to the decision-making problem in the form of a network is constructed. This includes not only the relationships between the groups of elements or within them, but also the feedbacks.

The Simple Multi-Attribute Rating Technique (SMART) method is a multi-stage method that identifies decision makers, opportunities, attributes relevant to a given decision making problem, the values and weights of individual attributes, the decision is made and the analyzes of its sensitivity is made [4].

The ELECTRE methods (Fr. ELimination Et Choix Traduisant la REalité) [3] have a wide range of applications in a variety of decision-making problems. They include a group of methods (e.g. ELECTRE I, Iv, IS, II, III, IV, TRI) designed to solve various problems of multi-criteria decision making support. The choice of a particular method depends on the one hand on the kind of problem we are dealing with and on the other on the type of data we have. The ELECTRE methods assume the axiom of limited comparability of variants, expressed by the recognition of four basic relations: I - equivalence, P - strong preference, Q - weak preference, and R - incomparability. The basic rule used in the ELECTRE methods is comparing each variant with all others. In this way it is checked whether one variant can be regarded as having an advantage over each of the others.

The PROMETHEE method (Preference Ranking Organization METHod for Enrichment Evaluations), like ELECTRE, represent a group of methods [2]. The PROMETHEE methods use information about the degree of preference of the given variant relative to the other variants and the information on the extent to which the other variants are more preferred to the given variant.

The ORESTE method has been developed for the situations where the alternatives are ordered according to each criterion and the criteria themselves are ranked according to their importance [12]. This method uses independent rankings for the criteria and alternatives to each criterion.

The REGIME method [12] is based on an overriding analysis, and can be seen as an orderly generalization of comparison methods in pairs such as compatibility analysis. The REGIME basis is CiL compliance coefficients defined as the sum of weights for criteria for which the alternative *ai* is at least as good as *al*. The purpose of this method is to determine the Cil - Cli difference sign. If the sign of this difference is positive then the alternative *ai* is preferred over the *al* and vice versa (when the sign is negative).

Selection of the method is extremely important from the point of view of the output information following the evaluation. It also depends on the nature of the input information held, the quantity and common dependencies, if any, and the information (purpose) one wants to obtain as the output. Therefore, the indicators and measures used for evaluating of the individual components of the vehicle operational evaluation are not exhaustive because they represent one variable. They do not reflect the relationships between the individual variables and the strength (size) of this relationship.

For the operational evaluation of the vehicles one can also use artificial intelligence methods, and, above all, artificial neural networks.

The purpose of this article is to present the possibility of using artificial neural networks for:

- support making decisions related to the operation of vehicles used in transport services related to the delivery of products for production,
- forecasting the quality and operational efficiency of the motor vehicles in the transport service system.

Applied were the following research methods: analysis (used to identify the area of artificial intelligence), descriptive modelling (used to formulate and describe collected information), mathematical modelling, using artificial neural networks (for operational evaluation of motor vehicles).

2. Aspects of operational evaluation of transport services

Transport services are invariably an essential element of the economy and social life, enabling them to function effectively. Socio-economic development generates the need for the people to move and transport loads. Lack of consistency between transport and manufacturing activity significantly undermines development opportunities. In addition, high competition in this segment has led to the fact that the lowest price has ceased to be a guarantee of market advantage. These considerations are one of many aspects of interest in evaluating transport services in terms of vehicle operation. The problem of quality in the face of dynamic market changes is becoming particularly important for such reasons as: ever-increasing customer expectations, minimizing the duration of the service, guaranteeing the highest efficiency of the service, or the safety of the vehicle and the goods being transported. One of the more important dimensions of the assessment of the transport service is the evaluation from the point of view of vehicle operation. This assessment is a complex problem, due to numerous criteria, described by attributes that are not measurable or difficult to measure. As a rule, the criteria are of heterogeneous character, which further complicates the credible assessment. This problem solves the use of multi-criteria decision-making methods based on, for example, heuristic methods or fuzzy sets theory. Unfortunately, these methods, due to their mathematical construction, are difficult to implement. For this reason, simple, practical tools are being sought for utilitarian benefits. In situations where there is a full knowledge of the rules and a small complexity of the problem, exact algorithms (e.g. linear models) are used. For partial or complete lack of knowledge of the rules and high complexity of the problem, neural networks are used. Because of its complexity and multi-facetedness, operational evaluation belongs to the area of artificial intelligence use.

There are many criteria for the classification of transport services. The relevant areas that influence the evaluation areas include: the subject of the carriage (e.g. passenger transport market, freight transport market), the mode of transport used, the area of operation (e.g. local, national, international market), economic strength of the operators (carrier's market or user's market), etc.

This article attempts to carry out an operational evaluation whose objects are motor vehicles.

The essential features of transport services include:

- complexity - the transport service consists of a very large number of elements and relationships between these elements,
- probability - all states and events can not be predicted,
- dynamics - the implementation of transport service gets constantly intervened in, both in time and space.

The basic conditions for the functioning of transport services include:

- economic and legal aspects – e.g. financial system, transport legislation,
- technical aspects - such as vehicles, infrastructure, transport and handling equipment, along with many operational aspects,
- organizational aspects – e.g. rules of cooperation between the carrier and the customer, carrier's working time.

The quality aspects of providing transport services constitute a separate group of research areas. It is possible to distinguish three quality categories [23]:

- postulated by the user, which sets out their requests and wishes regarding the way the transport service is performed,
- offered by carriers, that is the offered supply of transport services that can be realized with current knowledge, technology and organizational resources,
- implemented by service providers.

The issues of evaluating transport services are dealt with in different ways. Most often, this analysis is done with respect to the delivery time, safety and reliability of the service, and the safety and reliability of the vehicles themselves [5]. On the other hand, Neo and others have analyzed the quality of services provided by logistical operators and have given the accuracy of the information, the accuracy of the assembling process and the timely deliveries as the basic indicators of the evaluation [13].

Taking the above into consideration, the evaluation of the transport service can be made in many other aspects, i.e. costs incurred, risks, resources: human, information, or vehicles used. Depending on the aforementioned conditions and the nature of the service performed, these aspects are differently interpreted. Although for risk estimation, for example, it is recommended to carry out this process in a specific order: defining scope, identifying hazards and predefining consequences, estimating risk, verification, documenting, and updating analysis [16].

In general, the evaluation of a transport service can be presented as a following function:

$$R_n(t)=f(w_{n,1}(t), w_{n,2}(t),...., w_{n,k}(t)) \quad (1)$$

where:

$R_n(t)$ – evaluation of n-th transport service in time t,
$w_{n,k}(t)$ – evaluation of the k-th requirement of the n-th transport service in time t.

The aspects presented above can be described by different measures, depending from which point of view the evaluation is made. But reliable vehicles operated in the course of the transport service are essential. Their quality depends mainly on the proper operation, which is determined by reliability and readiness. In turn, the readiness of the vehicle consists of elements such as: resistance to damage, serviceability and ensuring operating means [15]. Therefore, for the purpose of this work, the operational evaluation of the n-th transport service is defined as a following function:

$$E_n(t) = f(w_{n,u}(t), w_{n,o}(t), w_{n,w}(t), w_{n,st}(t)) \quad (2)$$

where:

$E_n(t)$ – operational evaluation of the n-th transport service in time t,
$w_{n,u}(t)$ – evaluation of the resistance to damage requirement of the vehicle performing the n-th transport service in time t,
$w_{n,o}(t)$ – evaluation of the serviceability requirement of the vehicle performing the n-th transport service in time t,
$w_{n,w}(t)$ – evaluation of the age requirement of the vehicle performing the n-th transport service in time t,
$w_{n,st}(t)$ – evaluation of the technical condition of the vehicle performing the n-th transport service in time t.

In order to perform an operational assessment of transport services, due to its multidimensional nature, tools are needed to find the relationships between sets of variables at a high complexity of the problem. This tool can be artificial neural networks. So further in this article the modelling is shown, using this software, of the operational vehicles evaluation applied for assessment of transport services and results of the authors own studies.

3. Neural modelling

Research on the possibility and use of artificial neural networks for the operational evaluation of transport services was carried out based on the services provided by external carriers to a company that produces and markets dairy products. Evaluation concerned transport services carried out on the domestic market using motor vehicles.

Table. 1. Requirements description of the transport service operational assessment E_n

Symbol of the assessment requirement	Requirement description of the transport service operational assessment E_n
$w_{n,u}(t)$	vehicle resilience to defects - number of defects occurring per unit of time (e.g. week, month, etc.),
$w_{n,o}(t)$	vehicle serviceability - the amount of hours the vehicle is in service,
$w_{n,w}(t)$	age of the vehicle - affects other reliability characteristics therefore this requirement is included in vehicle quality measurements,
$w_{n,st}(t)$	technical condition of the vehicle - this characteristics stems, among the others, from the other reliability assessment indicators and is assessed organoleptically by an expert, markings adopted: very good technical condition of the vehicle (vg) satisfactory (sat), not satisfactory (nsat).

Source: author's own compilation.

Table 2. Parameterized assessment of requirement $w_{n,u}$ - vehicle's resilience to defects

No.	Vehicle's resilience to defects [number / month]	Parameterized quality assessment	Descriptive quality assessment
1	0	1	high level of quality
2	1	0,75	very good level of quality
3	2	0,5	good quality level
4	3	0,25	low quality level
5	4	0	unacceptable level of quality

Source: author's own compilation.

Table 3. Parameterized assessment of requirement $w_{n,o}$ - vehicle serviceability

No.	Vehicle serviceability [number of hours]	Parameterized quality assessment	Descriptive quality assessment
1	0	1	high level of quality
2	0-1	0,75	very good level of quality
3	2-5	0,5	good quality level
4	6-10	0,25	low quality level
5	>10	0	unacceptable level of quality

Source: author's own compilation.

Table 4. Parameterized assessment of requirement $w_{n,w}$ – age of the vehicle

No.	Age of the vehicle [in years]	Parameterized quality assessment	Descriptive quality assessment
1	0-5	1	high level of quality
2	6-12	0,5	good quality level
3	>12	0	unacceptable level of quality

Source: author's own compilation.

Table. 5. Parameterized assessment of requirement $w_{n,st}$ – technical condition of the vehicle

No.	Technical condition of the vehicle	Parameterized quality assessment	Descriptive quality assessment
1	very good	1	high level of quality
3	accept	0,5	good quality level
5	not accept	0	unacceptable level of quality

Source: author's own compilation.

Table 6. Sample data for teaching neural network

No.	Vehicle's resistance to defects	Vehicle serviceability	Vehicle's age	Vehicle's technical condition	Evaluation (weighted average)	Decision
	number / month	number of hours / month	years	v.good/accept./not accept.		
1	0	0	1	very good	1,00	Positive
2	1	1	3	very good	0,87	Positive
3	0	0	6	very good	0,88	Positive
4	1	1	6	very good	0,75	Positive
5	1	1	7	accept	0,63	Positive
6	0	0	7	accept	0,76	Positive
7	1	6	8	accept	0,56	Negative
8	0	0	8	accept	0,76	Positive
9	0	0	8	very good	0,88	Positive
10	1	6	11	not accept	0,44	Negative
11	0	0	11	not accept	0,64	Positive
12	2	8	12	accept	0,46	Negative
13	0	0	13	accept	0,64	Positive
14	1	2	13	not accept	0,36	Negative

Source: author's own compilation.

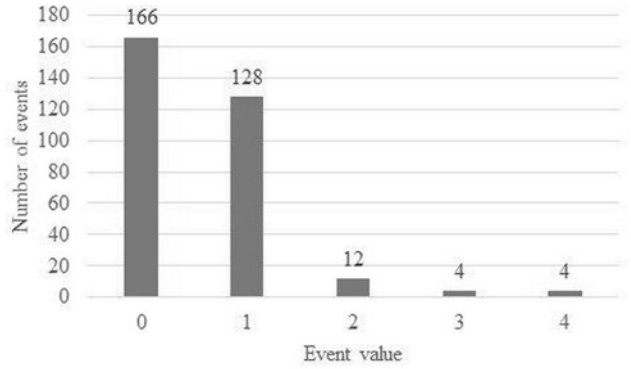


Fig. 1. Results of investigating the requirement $w_{n,u}$ – Vehicle's resistance to defects

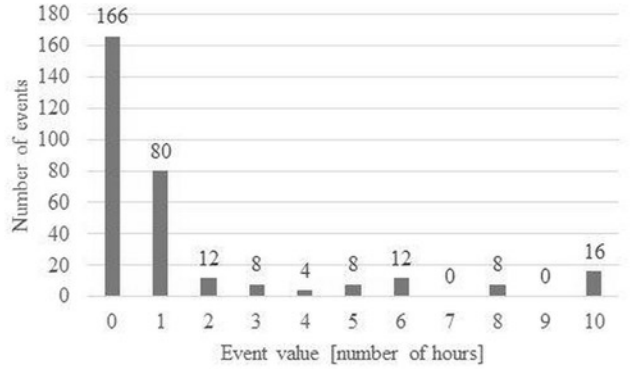


Fig. 2. Results of investigating the requirement $w_{n,o}$ – serviceability of the vehicle

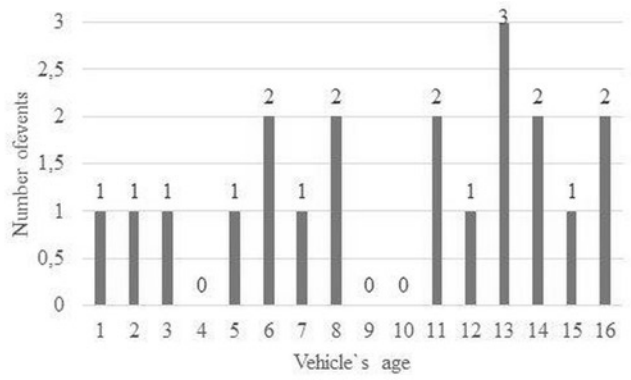


Fig. 3. Results of investigating the requirement $w_{n,w}$ – age of the vehicle

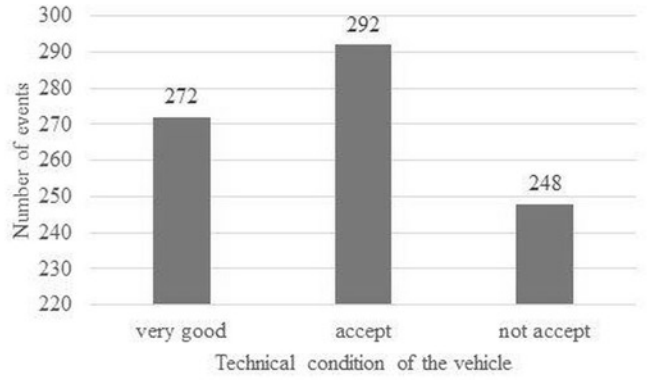


Fig. 4. Results of investigating the requirement $w_{n,st}$ – technical condition of the vehicle
Figures 1-4 source: author's own compilation.

In order to use the neural networks to carry out an in-service evaluation of transport services, a set of input signals, both quantitative (expressed in terms of numbers) and qualitative (expressed in words), was defined. The vehicles quality measurements describing the requirements of the operational evaluation of transport services are presented in the table 1.

For the purpose to assess the quality of the transport service $E_n(t)$ according to the requirements $w_{n,u}(t)$, $w_{n,o}(t)$, $w_{n,w}(t)$, $w_{n,st}(t)$, they have been parameterized as per arrangements given in tables 2-5.

To evaluate individual requirements, the experts assigned weights to the characteristics on the scale (0-10), where 0 is not significant, 10 is very important. The following weight values are assigned:

- $w_{n,u}$ – 8,
- $w_{n,o}$ – 3,
- $w_{n,w}$ – 5,
- $w_{n,st}$ – 5.

With so defined operational assessment requirements, vehicle users and experts assessed the individual requirements. The data collected from the operation was from the implementation of 812 transport services completed in the last 5 years, supplying products for production. This data provided the starting point for the studies (table 6). Based on the weighted average of the individual requirements, the decisions were made on the operational evaluation of the quality of transport services (positive or negative). The quality level considered satisfactory was assumed at 0.6 and above.

Below is a summary of the structure of the results obtained, according to the individual requirements.

Based on the above data, 540 positive and 272 negative evaluations were received.

Of the many types of neural networks and many of their teaching algorithms, further studies have used the Multilayer Perceptron and teaching algorithms: the fastest drop method, the conjugate gradient method; BFGS method (Broyden-Fletcher-Goldfarb-Shanno). The neural network used belongs to the following groups:

- the so-called, supervised networks, where the teaching process takes place under the supervision of the teacher (among the outgoing signals there is a master signal),
- unidirectional networks where the flow of signals (information) takes place in one direction (from the input to the output of the neural network).

Using the Statistica 12 computer program, transportation services have been evaluated using predefined vehicle quality measures.

The following signals were thus identified:

- input quantitative ones : $w_{n,u}(t)$, $w_{n,o}(t)$, $w_{n,w}(t)$,
- input qualitative ones: $w_{n,st}(t)$,
- output quantitative ones: $E_n(t)$.

With the input data indicated, the size of the sets was defined. It was stated that:

- 80 % - of the data will be the teaching set used to modify the weights,
- 10 % - a test set for ongoing monitoring of the teaching process,
- 10 % - validation set for network quality assessment after completing teaching process.

Then the basic parameters of the network were defined, i.e.:

- network type (multi-layer perceptron (MLP)
- minimum number of hidden neurons
- maximum number of hidden neurons
- number of teaching networks,
- the number of networks retained,
- hidden neuron activation function,
- activation function of output neurons,

• weight reduction values for the hidden layer and the output layer.

Once the data and network parameters were defined, neural network teaching was performed using the collected data. The sample results of this process are shown in the table 7.

With such defined requirements and having conducted the teaching process, the structure of the best network became MLP 6-3-1, which means 6 neurons in the input layer, 3 neurons in the hidden layer and 1 neuron in the output layer (figure 5).

The teaching quality of MLP 6-3-1 network was rated at 99,6% probability of indicating a correct (response), while testing quality at a 99,7%, which means that all tests in this set have been properly assigned and the quality of the validation has been determined to be at 99,4%. The best teaching algorithm turned out to be BFGS 148 (number 148 means the number of epochs the network needed to carry out the teaching process and finding the best network, with the smallest error).

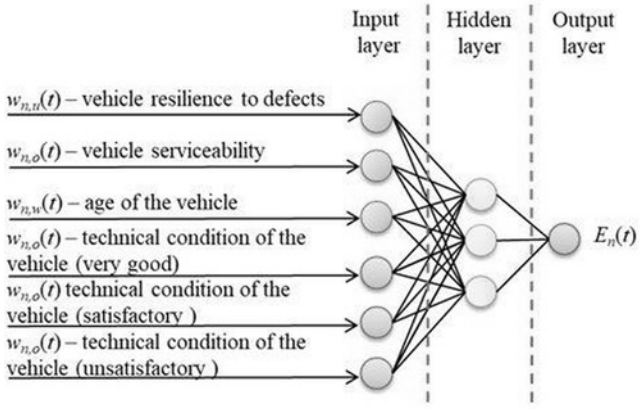


Fig. 5. Structure of the investigated MLP 6-3-1 network
Source: author's own compilation

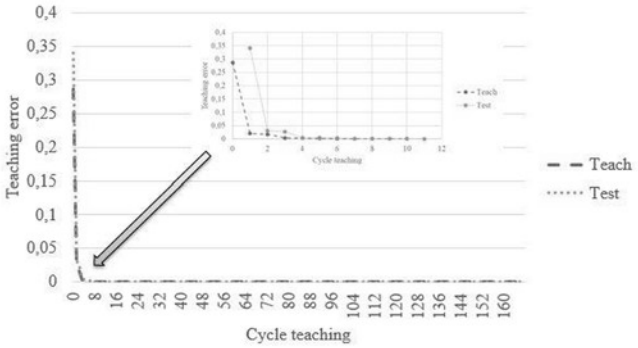


Fig. 6. Neural network MLP 6-3-1 teaching diagram
Source: author's own compilation.

Table 7. Sample results of teaching neural network

No.	Network Name	Teaching Quality	Testing Quality	Validation Quality	Teaching Algorithm	Error Function	Hidden Activation	Output Activation
1	MLP 6-5-1	0,988	0,992	0,989	BFGS 55	SOS	Logistic	Logistic
2	MLP 6-4-1	0,988	0,992	0,989	BFGS 98	SOS	Logistic	Tanh
3	MLP 6-6-1	0,984	0,990	0,991	BFGS 10	SOS	Linear	Linear
4	MLP 6-7-1	0,984	0,991	0,990	BFGS 8	SOS	Linear	Linear
5	MLP 6-8-1	0,983	0,991	0,991	BFGS 10	SOS	Linear	Linear
6	MLP 6-4-1	0,983	0,992	0,991	Fastest Drop 30	SOS	Tanh	Linear
7	MLP 6-4-1	0,987	0,993	0,989	BFGS 47	SOS	Logistic	Exponential
8	MLP 6-8-1	0,987	0,992	0,987	BFGS 48	SOS	Logistic	Logistic
9	MLP 6-7-1	0,990	0,993	0,989	BFGS 63	SOS	Logistic	Linear
10	MLP 6-3-1	0,996	0,997	0,994	BFGS 148	SOS	Tanh	Linear
11	MLP 6-8-1	0,983	0,991	0,991	BFGS 9	SOS	Linear	Linear
12	MLP 6-4-1	0,985	0,992	0,988	Conjugate gradients 27	SOS	Tanh	Linear
13	MLP 6-6-1	0,977	0,984	0,990	BFGS 12	SOS	Logistic	Sinus
14	MLP 6-4-1	0,981	0,984	0,990	BFGS 23	SOS	Sinus	Sinus
15	MLP 6-6-1	0,981	0,988	0,990	BFGS 23	SOS	Sinus	Sinus
16	MLP 6-10-1	0,996	0,995	0,997	BFGS 87	SOS	Tanh	Linear
17	MLP 6-5-1	0,991	0,995	0,990	BFGS 86	SOS	Tanh	Tanh
18	MLP 6-8-1	0,985	0,992	0,988	BFGS 16	SOS	Exponential	Logistic
19	MLP 6-14-1	0,991	0,993	0,990	BFGS 47	SOS	Tanh	Exponential
20	MLP 6-1-1	0,976	0,983	0,990	BFGS 16	SOS	Linear	Sinus
21	MLP 6-1-1	0,985	0,992	0,987	BFGS 39	SOS	Exponential	Exponential

Source: author's own compilation.

4. Verification of the selected neural network MLP 6-3-1

The prove of the positive result of teaching neural network is provided by the teaching curve which shows that the best network structure was found in 148 epoch where the share of incorrect answers is below 1%, and the error was estimated at 0,0002.

The errors matrix are shown in the Table 8. It indicates exactly how many cases of evaluation have been qualified by the network as a positive (above 0,6) or negative (below 0,6) evaluation. The table shows that for 432 positive evaluations, the network correctly assigned 408 indications, while correctly indicating all negative evaluations.

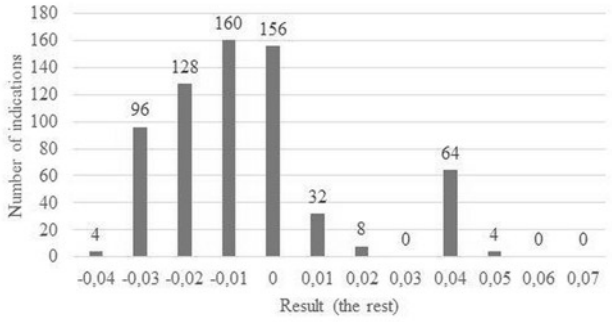


Fig. 7. Distribution of residues neural network MLP 6-3-1
Source: author's own compilation.

Table 8. The errors matrix of the neural network MLP 6-3-1

No.		Negative Decision	Positive Decision	Decision - all
1	Total	220	432	652
2	Correct	220	408	628
3	Incorrect	0	24	24
4	Correct (%)	100%	94%	96%
5	Incorrect (%)	0%	6%	4%

Source: author's own compilation.

Table 9. Predictions for new data based on MLP 6-3-1 neural network

No.	Vehicle resilience to defects	Vehicle serviceability	Age of the vehicle	Technical condition of the vehicle	Evaluation
1	0	0	3	Very good	0,99
2	1	5	3	Very good	0,863
3	1	2	5	Accept.	0,726
4	2	10	7	Accept.	0,552
5	1	5	10	Not accept.	0,458

Another important feature of the study of the neural network is the distribution of residues shown in figure 7, i.e. the differences between the output variable and its prediction.

From the histogram can be seen that the residues are normally distributed around zero with an emphasis on negative values. The vast majority of evaluations were made with an error level -0,03-0.

The last stage of verification of the neural network are the predictions for new inputs (for attempts that have not yet appeared in any collection). In order to obtain new operational evaluations of the quality of transport service the values of all input signals to the neural network were supplemented based on which the output signal (i.e. operational evaluation (quantitative) of transport service) has been generated. After entering data into the network the final results were obtained. Table 9 contains predictions for new data based on MLP 6-3-1 neural network.

The obtained results indicate the possibility of using a single-layer neural network to perform an in-service evaluation of the quality of transport services. Both, the number and type of data (quantitative or qualitative), do not affect high performance at 98% - 99% of efficiency. Based on data from the past, a neural network allows making decisions, generating assessment of the current or future operation.

5. Conclusion

There are many different methods and models (mostly multi-criteria), in the subject literature, for the purpose of evaluating operation of vehicles and technical systems. They are characterized in the introduction to this article. They are commonly used in practice. The authors' experiences and the conducted analysis of the situation indicate that neural networks are not yet widely used for the operational evaluation of vehicles to be used in transport services. Neural networks are used in operation but in other areas [7, 19, 20, 21].

The results of the investigations carried out obtained at the company manufacturing dairy products and at the transport companies cooperating with this company indicate the possibility of using artificial neural networks to evaluate vehicles used in the delivery of transport services, using operational quality measures. In the case of negative assessments, corrective action can be taken without delay. Based on this it is possible to forecast future predictions with the data from current operation.

The proposed operational quality measures resulted from the needs of the surveyed company. In practice, they can be selected in different ways, depending on the purpose of the analysis.

Neural networks have proved to be useful as a tool for:

- supporting the decision making related to the use of vehicles used in transport services for the delivery of goods for production,
- forecasting the quality and efficiency of the operation of motor vehicles in the transport service system.

So the purpose of the research has been achieved.

The use of neural networks in operation can be broader. For example, to assess not only operational risks, but also related to the reliability and safety of vehicles (not just cars) and other machinery and to the safety of the transport services themselves.

The computational example presented in the article thus reflects the essence of using artificial neural networks to evaluate operational issues. This is one possible starting point for further research into the use of neural networks in this area.

References

1. Adamus W, Gręda A. Wspomaganie decyzji wielokryterialnych w rozwiązywaniu wybranych problemów organizacyjnych i menedżerskich. Badania operacyjne i decyzje 2005; 2.
2. Brans J P, Mareschal P. Promethee methods, [in:] Figueira J., Greco S., Ehrgott M. (ed.), Multiple Criteria Decision Analysis: State of the Art Surveys, Springer Verlag, Boston, Dordrecht, London, 2005, https://doi.org/10.1007/0-387-23081-5_5.
3. Figueira J, Mousseau V, Roy B. ELECTRE methods, [in:] J. Figueira, S. Greco, and M. Ehrgott (eds.), Multiple Criteria Decision Analysis: State of the Art Surveys, Springer Verlag, Boston, Dordrecht, London, 2005, https://doi.org/10.1007/0-387-23081-5_4.
4. Goodwin P, Wright G, Analiza decyzji, Wolters Kluwer Polska, Warszawa 2011.
5. Grigoroudis E, Siskos Y. A survey customer satisfaction barometers: Some results from the transportation-communications sector. European Journal of Operational Research 2004; 152: 334.353.
6. Jacyna-Gołda I, Lewczuk K. The method of estimating dependability of supply chain elements on the base of technical and organizational redundancy of process. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (3): 382-392, https://doi.org/10.17531/ein.2017.3.9.
7. Karpenko M, Sepehri A. Neural network classifiers applied to condition monitoring of a pneumatic process valve actuator. Engineering Applications of Artificial Intelligence 2002; 15 (3): 273-283, https://doi.org/10.1016/S0952-1976(02)00068-4.
8. Kaźmierczak J. Eksploatacja systemów technicznych. Gliwice: Wydawnictwo Politechniki Śląskiej, 2000.
9. Knopik L, Migawa K. Multi-state model of maintenance policy. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (1): 125-130, https://doi.org/10.17531/ein.2018.1.16.
10. Kornacki A, Sokołowska E. The estimation of smooth operation time until failure with the application of the Akaike Information Criterion (AIC). Eksploatacja i Niezawodnosc - Maintenance and Reliability 2010; 1 (45): 69-76.
11. Loska A. Exploitation assessment of selected technical objects using taxonomic methods, Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15 (1): 1-8.
12. Martel J M. Other Outranking Methods, [w:] Multiple Criteria Decision Analysis: State of the Art Surveys, J. Figueira, S. Greco, and M. Ehrgott, editors, Springer Verlag, Boston, Dordrecht, London, 2005.
13. Neo H-Y, Xie M, Tsui K._Service quality analysis: case study of a 3PL company. International Journal of Logistics Systems and Management 2004; 1(1): 64-80, https://doi.org/10.1504/IJLSM.2004.005539.
14. Niebel W. B. Engineering Maintenance Management. Second edition. New York: Marcel Dekker Inc., 1994.
15. PN-EN 15341:2007 - Maintenance - Maintenance Key Performance Indicators.
16. PN-EN 60300-1:2015 Dependability management - Part 1: Guidance for management and application.
17. Saaty T. Decision making for Leaders The Analytic Hierarchy Process for decisions in a complex world, University of Pittsburgh, RWS Publications, Pittsburgh 2001.
18. Saaty T. The Analytic Hierarchy Process. RWS Publications. Pittsburgh 1998.
19. Samanta B. Gear Fault Detection Using Artificial Neural Network and Support Vector Machines with Genetic Algorithms. Mechanical Systems and Signal Processing 2004; 18: 625-644, https://doi.org/10.1016/S0888-3270(03)00020-7.
20. Samanta B, Al-Balushi K R. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. Mechanical Systems and Signal Processing 2003; 17 (2): 317-328, https://doi.org/10.1006/mssp.2001.1462.
21. Samanta B, Al-Balushi K R, Al-Araimi S A. Use of genetic algorithm and artificial neural network form gear condition diagnostics. Proc. of COMADEM, University of Manchester 2001: 449-456.
22. Szudrowicz M. Layered composite increasing the resistance of patrol and intervention vehicles to the impact of improvised explosive devices (iEd) from below. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (1): 9-15, https://doi.org/10.17531/ein.2018.1.2.
23. Świderski A. Modelowanie oceny jakości usług transportowych, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, 2011.
24. Siskos Y, Grigoroudis E, Matsatsinis N F. Multiple criteria decision analysis: State of the art surveys, Springer New York, 2005.

Andrzej ŚWIDERSKI
Motor Transport Institute
ul. Jagiellońska 80, 03-301 Warsaw, Poland

Arkadiusz JÓŹWIAK
Military University of Technology, Faculty of Logistics
ul. gen. Witolda Urbanowicza 2, 00-908 Warsaw, Poland

Roland JACHIMOWSKI
Warsaw University of Technology, Faculty of Transport
Koszykowa 75, 00-662 Warsaw, Poland

Emails: andrzej.swiderski@its.waw.pl, arkadiusz.jozwiak@wat.edu.pl, rjach@wt.pw.edu.pl

Article citation info:
DONG F, LIU Z, WU Y, HAO J. A multi-stage risk-adjusted control chart for monitoring and early-warning of products sold with two-dimensional warranty. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 300–307, http://dx.doi.org/10.17531/ein.2018.2.17.

Fangqi DONG
Zixian LIU
Yujie WU
Jianhong HAO

A MULTI-STAGE RISK-ADJUSTED CONTROL CHART FOR MONITORING AND EARLY-WARNINGOF PRODUCTS SOLD WITH TWO-DIMENSIONAL WARRANTY

KARTA KONTROLNA DO WIELOETAPOWEGO MONITOROWANIA PRODUKTÓW SPRZEDAWANYCH Z GWARANCJĄ DWUWYMIAROWĄ, Z KOREKTĄ RYZYKA I WCZESNE OSTRZEGANIE O WADACH PRODUKCYJNYCH NA PODSTAWIE DANYCH Z REKLAMACJI

Warranty claims data contain valuable information about the quality and reliability of products. The monitoring and early-warning of warranty claims data are of great significance to the manufacturer by identifying and solving the emerging quality or reliability problem as soon as possible. However, though it has been used widely in the automobile industry, there are no studies that have been carried out on the monitoring and early-warning of claims data for products sold with two-dimensional warranty. In order to fill this gap, fitting the two-dimensional warranty claims data with accelerated failure model (AFT), a multi-stage risk-adjusted control chart is proposed by this paper, for which a reasonable product sales tracking time and a monitoring time are suggested to reduce the influence of sales delay and fluctuating claim rates. Comparing with traditional Cumulative Sum Control Chart (CUSUM), the applicability and availability of the proposed model are demonstrated in the final.

Keywords: *two-dimensional product warranty, claims data, monitoring and early-warning, multi-stage control chart, accelerated failure model, risk adjustment.*

Roszczenia gwarancyjne stanowią cenne źródło informacji na temat jakości i niezawodności produktów. Monitorowanie danych dotyczących roszczeń gwarancyjnych i wczesne ostrzeganie w oparciu o te dane ma wielkie znaczenie dla producenta, ponieważ pozwala rozpoznawać i rozwiązywać pojawiające się problemy związane z niezawodnością w jak najkrótszym czasie. Chociaż ten rodzaj monitorowania i wczesnego ostrzegania jest szeroko stosowany w przemyśle motoryzacyjnym, nie przeprowadzono dotąd żadnych badań na temat tych procesów w odniesieniu do produktów sprzedawanych z gwarancją dwuwymiarową. W celu wypełnienia tej luki, dane o reklamacjach składanych na podstawie gwarancji dwuwymiarowych dopasowano modelem uszkodzeń przyspieszonych (accelerated failure model, AFT), a następnie przedstawiono koncepcję karty kontrolnej monitorowania wieloetapowego z korektą ryzyka, dla której zaproponowano odpowiedni czas śledzenia sprzedaży produktu i czas monitorowania, mając na uwadze zmniejszenie wpływu opóźnień w sprzedaży i wahań liczby roszczeń zgłaszanych z tytułu gwarancji. Możliwości zastosowania i dostępność proponowanego modelu porównano z tradycyjną kartą sum skumulowanych.

Słowa kluczowe: *dwuwymiarowa gwarancja na produkt, dane o roszczeniach z tytułu gwarancji, monitorowanie i wczesne ostrzeganie, karta kontrolna procesu wieloetapowego, model przyspieszonego uszkodzenia, korekta ryzyka.*

1. Introduction

In today's competitive market, offering attractive warranty service has been used by many manufacturers to capture more market share and customers' satisfaction. Warranty is a contractual agreement between the manufacturer and the customer, which specifies the manufacturer's obligation in the event that the product is unable to perform satisfactorily when properly used [1]. Based on the number of variables used to define warranty coverage, warranty policies can be broadly divided into two categories: one-dimensional warranty and two-dimensional warranty [2]. One-dimensional warranty is usually defined on the basis of age or usage, while two-dimensional warranty takes age and usage or the potential interaction between them into consideration. In practice, two-dimensional warranty has been widely applied in automobile industry. For example, a new automobile is usually sold with a two-dimensional warranty, offering free repair for 3 years or 60000 km, whichever occurs first.

With the intensification of market competition, the percentage of warranty cost to the profit of manufacturers is gradually increasing. For the US automotive industry, manufacturers spend roughly $10 billion–$13 billion per year on warranty claims, consuming roughly 1%–5.2% of the product profit [3]. Among various factors affecting warranty cost, product reliability and quality control have attracted significant attention from practitioners and academics. So far, reliability design, quality control management, and many other reliability technologies have been widely used in the design and development stages of products, which have resulted in great contributions to improve product reliability and quality [4]. However, many reliability or quality issues still occur during the product warranty period, such as unknown failure modes, unanticipated changes in operating environ-

ments, unknown changes in raw materials, etc., which may result in huge warranty costs. As warranty data contain valuable information about the quality and reliability of products, the use of appropriate statistical detection rules in a warranty database has the potential to identify or warn serious reliability problems long before they would otherwise be discovered [5]. Detecting these problems, a month or even a week earlier, can effectively reduce tangible or intangible costs for the manufacturers, especially for products sold with two-dimensional warranty, such as automobile industry, the warranty cost of which is relatively expensive.

The research about warranty data analysis have been classified into five different directions by Wu [5], including identifying early warnings of abnormalities in their products, providing useful information for product modification and improvement; estimating and explaining the costs of warranty claims; predicting failure claims and warranty costs, and estimating product reliability [4-7]. With the aim to provide early indications of unexpected quality or reliability problems for the manufacturer, the monitoring and early-warning of warranty claims data have become the most important aspects for the research about warranty data.

In general, the model for monitoring and early-warning of warranty claims data can be divided into two stages: first, it is to analyze and fit warranty claims data with parameter model or non-parameter model; next, the control chart for monitoring warranty claims data is proposed to balance Type-1and Type-2 error in the monitoring models [8].As far as we know, most of the monitoring and early warning research have only considered the product sold with one-dimensional warranty, for which the impact of correlation time on the claim rate is conspicuous. Therefore, in the first stage of modeling, claims data are always hierarchically aggregated according to product production time, product sales time and product usage time [3,4,9], the reason/objective for doing this is to analyze and model the influence of product correlation time on product claim rate clearly,then the detecting and tracing analysis for abnormal conditions can become more easily. However, for product sold with two-dimensional warranty, the warranty claim rate is assumed to be affected by the time and usage of a product simultaneously, so fitting warranty claims data only with correlation time cannot meet the need for monitoring of products sold with two-dimensional warranty, which deserves further research.

In the second stage, for product sold with one-dimensional warranty, a variety of statistical techniques, theories and methods, such as Shewhart control chart, Cumulative Sum Control Chart (CUSUM), change-point theory, artificial intelligence algorithms, and many other tools have been proposed for monitoring and early warning of warranty claims data [3,4,9,10,15]. For example, based on change-point theory, a monitoring model is provided by Karim [9] to detect whether the product reliability has been changed and record the time and the form it happened. Classifying and fitting warranty claims data by production time, sales time and usage time, a multi-stage control chart is proposed by Wu and Meek [4] to monitor claims data month by month, which has laid the foundation for the research of warranty claims data monitoring. To improve the monitoring capability of the above models, the traditional Cumulative Sum Control Chart (CUSUM) is suggested by Lawless [10] to monitor warranty claims data subsequently. At present, monitoring warranty claims data according to the correlation time independently can lead to many control charts working at the same time, which may result in poor robustness and high false alarm rates. Furthermore, when the quantity of products monitored at different stages or types is small, the over classification of warranty claims data may result in many inefficient monitors, which have also restricted the applicability of the control chart.

No studies have been focused on the monitoring and early-warning of products sold with two-dimensional warranty ,however, the monitoring of one-dimensional warranty claims data with different covariates in the previous studies have built the foundation for it. For instance, considering supply chain quality information as key covariates, a multi-stage monitoring model with cox proportional hazard (PH) model is proposed by Zhou [3]to monitor the warranty claims data with a relatively high claim rate. Taking into account some qualitative factors, such as automobile type, warranty service area, seasonal factors and so on, a warning threshold for the monitoring of warranty claims data is obtained with Artificial Neural Networks (ANNs) and analytic hierarchy process (AHP) by Lee [10]. Subsequently, this model has been extended by Na [11] to monitor warranty claims data with both qualitative and quantitative influencing factors. For product under uncertain service environment, a fuzzy feedback control method is proposed by Lee [13] to adjust the warning threshold according to the environmental implication.

The pertinent literatures have been reviewed above. For monitoring and early-warning of products sold with two-dimensional warranty, the influence of product usage rate on warranty claims data has to be considered. This is the main difference between monitoring one-dimensional and two-dimensional warranty claims data. In this study, on the condition of fitting two-dimensional warranty claims data with Accelerated failure time model (AFT), a multi-stage risk-adjusted control chart is provided for the monitoring and early-warning of two-dimensional warranty claims data. Firstly, considering the information can be used at different stages, risk adjustment has been carried out to allocate the false rate probability. Secondly, to reduce the effect of sales delay and fluctuating claim rates on the monitoring capability, a reasonable limit of product sales tracking time and monitoring time is suggested in this study. Finally, comparing with traditional CUSUM control chart, the availability and applicability of the proposed control chart are demonstrated by a case study.

2. Modeling two-dimensional warranty claims data

2.1. Introduction of warranty data

Warranty data are composed of claims data and supplement data [5]. Claims data are the data collected during the servicing of claims under warranty, and supplement data are additional data (such as production and marketing related, warranty cost, items with no claims, etc.) that are needed for effective warranty management. Both claims data and supplement data are usually collected according to the fail-

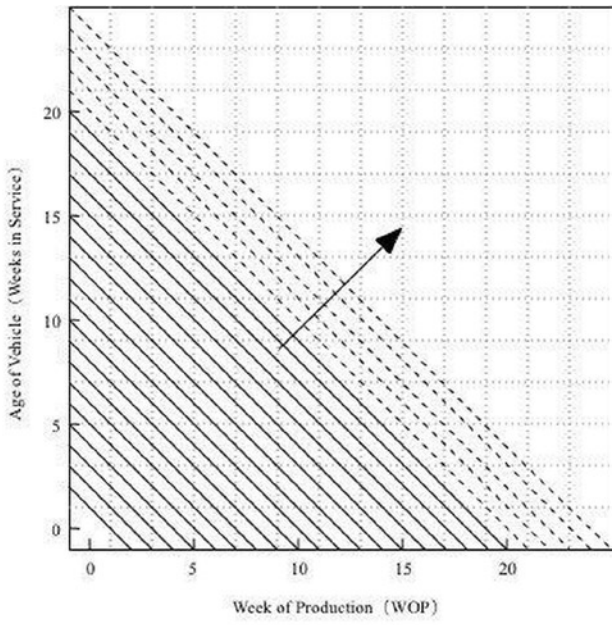


Fig .1.Warranty claim quantity for different monitoring stages

ure mode of products, which can provide valuable information about product quality and field reliability.

As we all know, the production and sales time of products are continuous for the manufacturer, and it is necessary to record and monitor the warranty claims data at different stages continuously, such as weekly or monthly, so as to detect the abnormal conditions of claims data as soon as possible. For a specified monitoring period or stage, the monitoring index of warranty claims data can be the whole claim quantity of products, the whole warranty cost of products, or the claim quantity for a specified failure mode, and so on. All of them can provide useful information to the manufacturer. In this paper, we choose the claim quantity for a specified failure mode as the monitoring index, which can also be extended by other options. As shown in Fig 1, the claim quantity of products that can be collected and monitored at different stages is changing gradually.

2.2. Warranty claims data analysis

For products sold with two-dimensional warranty, the claim rates of them are also affected by product correlation time, such as production time, sales time and service time. So, it is still necessary to classify warranty claims data according to correlation time. Considering one week as the smallest unit of time, it is supposed that products are produced and sold at the first week or stage, as shown in Fig 2. N_i is the quantity of products produced in stage i; N_{ij} is the quantity of products produced in stage i and sold in stage j; R_{ijk} is the quantity of products produced in stage i, sold in stage j, and claimed in stage k, where j and k represent the relative sales stage and service stage, respectively. For example, R_{111} represents the claim quantity of products that occurred during the first stage.

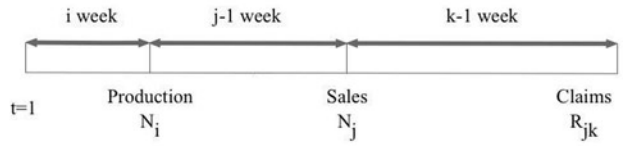


Fig. 2. Production, sale, and services schedule

In addition to the correlation time, the usage rate is also a main factor influencing the claim rate ofa product sold with two-dimensional warranty. In order to model the claim rate of a product in terms of its age and usage, three different approaches have been proposed for modelling claims data under two-dimensional warranty, including marginal approach, bivariate failure distribution approach and composite scale approach [14]. Readers can refer to [21] for a brief review of the recent publications on these approaches. For the current study, Accelerated failure time (AFT) and proportional hazard (PH) models, as well-known variations of the marginal approaches, are widely used to model the effect of customer usage rate on the claim rate of products [25]. In most recent researches, Iskandar and Jack [22], Jack et al. [23], and Baik and Murthy [24] have used AFT model to investigate the effect of usage rate on the claim rate of products, which would also be followed in this study.

Assume X_d with cumulative distribution function $F_d(x:\alpha_0,\beta)$, is the time to the first failure of the product under the nominal usage rate r_0, α_0 and β are the scale parameter and shape parameter of the distribution model of X_d, respectively.In the marginal approach, it is assumed that the usage rate of a given customer over the warranty period is constant but it varies randomly from customer to customer. Suppose r is the usage rate (random variable), and $g(r)$, $G(r)$ are its probability density function (PDF) and cumulative distribution function(CDF) respectively. According to the AFT model, for a specified usage rate r, the cumulative distribution function of the time to the first claim of the product X_r,will be:

$$\mathrm{F}\left(\mathrm{x}:\alpha\left(\mathrm{r}\right),\beta\right)=F_D\left(x:\alpha_0\left(\frac{r_0}{r}\right)^{\gamma},\beta\right) \tag{1}$$

Where $\gamma\geq 1$ is the accelerating failure factor. The nominal usage rate r_0 is also defined as experimental utilization rate in the model of AFT[24], which is often set as an average usage level in practice, then:

$$r_0=\int_0^{+\infty} rg(r)dr \tag{2}$$

Suppose $z_D(x)$ and $z_r(x)$ are the product hazard rate functions for the nominal usage rate r_0, and a random usage rate r, respectively, they can be calculated by:

$$z_D(x)=\frac{dF\left(\mathrm{x}:\alpha,\beta\right)/dx}{1-F\left(\mathrm{x}:\alpha,\beta\right)}$$
$$z_r(x)=\frac{dF\left(\mathrm{x}:\alpha\left(\mathrm{r}\right),\beta\right)/dx}{1-F\left(\mathrm{x}:\alpha\left(\mathrm{r}\right),\beta\right)} \tag{3}$$

3. The multi-stage risk-adjusted control chart

In this section, a multi-stage risk-adjusted control chart is provided to monitor the two-dimensional warranty claims data. Firstly, the proposed multi-stage control chart means monitoring the warranty claims data at different stages independently and continuously, such as weekly or monthly. The actual claim quantity, the theoretical distribution function of warranty claims data at different stages are analyzed elaborately. Secondly, considering the information available at different monitoring stages, risk adjustment has been considered to allocate the overall false rate of the multi-stage control chart, which can balance quick detection and the overall probability of detection for potential reliability problem. Then, the monitoring thresholds of the proposed control chart are calculated subsequently. Finally, in order to reduce the influence of sales delay and fluctuating claim rates on the monitoring capability of the control chart, reasonable product sales tracking time and monitoring time are suggested in this study.

3.1. Monitoring index

As mentioned above, the production and sales of products are continuous to the manufacturer. Monitoring warranty claims data independently according to different production time, sales time and service time can result in many issues, such as high false alarm rate, poor robustness, inefficient monitor, and so on. In this study, for a specified monitoring stage, the claim quantity of products with different production and sales time are monitored as an overall monitoring index, which can avoid the problems mentioned above largely. Suppose M is the prespecified number of stages for monitoring, taking monitoring stage l as an example, the actual warranty claim quantity of products monitored at stage l is:

$$\mathrm{o}_l=\sum_{i+j+k-2=l} R_{ijk},\forall i+(j-1)+(k-1)\leq M \tag{4}$$

As the monitoring and early warning of warranty claims often start at the sales time of products, and stop long before the deadline of warranty time limit, there is no need to consider the extreme case that the product under monitoring will exceed the warranty period because of product usage.

Subsequently, the theoretical distribution function of warranty claims data at different stages need to be estimated. In order to do this, the maintenance strategies applied for product claims during product warranty period have to be considered. According to degree of restorability of the product, Pham and Wang [17] classified the maintenance actions during warranty into three main categories, including minimal repair, imperfect repair and perfect repair. Minimal repair means restoring the product or system to its as-good-as old operation condition [18-20]. For a complex system as automobile, the claim of it may be caused by failures of one or few components, and simply repair or replacing these items would not significantly improve or reduce the system reliability. Therefore, the application of minimal repair for products such as automobile is an appropriate assumption. At present, minimal repair strategy has been widely used in the automotive industry, electronic products, and other complex systems.

Assuming minimal repairs with negligible repair times applied for warranty claims, then, the claim quantity up to the time t will constitute a non-homogenous Poisson process (NHPP), the intensity function of which is the same as the hazard function of the time to the first failure of the product. Thus, as given in eq. (3), for a given probability density function (PDF) of product usage rate, $g(r)$,the expected claim quantity at different service stage are:

$$\gamma_k^0 = \int_{r_{min}}^{r_{max}} g(r) \int_{k-1}^{k} z_r(x) dx dr, k = 1,2\ldots \tag{5}$$

Where r_{min} and r_{max} are the minimum and maximum usage rates for customer respectively. Then, the theoretical claim quantity at different monitoring stages will also follow Non-Homogeneous Poisson Process (NHPP) with different hazard rate functions. According to this, we can calculate the quantity of theoretical product warranty claims e_l in monitoring stage l, which is:

$$e_l = \sum_{i+j+k-2=l} N_{ij} * \gamma_k^0, \ \forall i + (j-1) + (k-1) \le M \tag{6}$$

As any failures occurring during the warranty period will result in one claim and maintenance with minimal repair, the claim quantity at different monitoring stages will follow NHPP independently. Therefore, the method proposed in this study can monitor the claim quantity of products at different monitoring stages independently.

3.2. Risk adjustment for the multi-stage control chart

For the multi-stage control chart proposed in this study, the monitoring at different stages are working independently. However, as the valuable information collected and monitored at different stages are not the same, it is unreasonable to allocate the false alarm probability and power of detection identically. To deal with this problem, risk adjustment has been considered, for which the overall false alarm rate of the multi-stage control chart is allocated according to the available information at different monitoring stages.

Define α_l as the false alarm rate at the l monitoring stage, $l \in [1, M]$. As the monitoring of claims data at different stages is independent, the overall false alarm rate for the multi-stage control chart up to stage M is given by:

$$\alpha = 1 - \prod_{l=1}^{M} (1 - \alpha_l) \tag{7}$$

For a given overall false alarm rate α, increasing M, the prespecified number of monitoring stages, will lead to the decrease of the false alarm rate allocated at different monitoring stages. But meanwhile, the early-warning ability, which means the detection ability for abnormal conditions of the control chart, has been increased. Therefore, M should be set properly to balance the false alarm rate and the early-warning ability.

The available information at different stages has an important influence on the balance between the false alarm rate and the early-warning ability. In this paper, the expected warranty claims e_l, is used to represent the information available at stage l, then the false alarm rate is allocated proportionally to the expected claims at different stages, that is:

$$\alpha_l = \eta * \sum_{i+j+k-2=l} N_{ij} * \gamma_k^0 = \eta * e_l \tag{8}$$

η is a constant to ensure that $1 - \prod_{L=1}^{M} (1 - \alpha_L) = \alpha$, which can be calculated by interpolation method. Risk adjustment can not only effectively reduce the high false alarm rate caused by the limited information in the early period of the monitoring, but also enhance the robustness of the proposed monitoring model. According to the allocation of false alarm probability, the monitoring threshold C_l of the monitoring stage l can be calculated by distribution parameter $\lambda = e_l$ and the false alarm rate α_l, which is given by:

$$P(O_l > C_l | O_l \sim P(e_l)) = \alpha_l \tag{9}$$

In other words, the monitoring threshold C_l is the α_l upper quartile for the NHPP distribution, with parameter $\lambda = e_l$ .Because of the independence of the monitoring at different stages, the overall false alarm probability for the proposed model up to stage l is given by:

$$\alpha_{2l} = 1 - \prod_{k=1}^{l} (1 - p(o_l > c_l)) \tag{10}$$

It should be noted that α_l is the false alarm rate at stage l, while α_{2l} is the overall false alarm rate up to stage l, which may be easily confused. The response time of abnormal conditions of the proposed control chart cannot be accurately predicted. In this study, $T = l$ is define as the early warning response time, up to when the overall alarm probability for the abnormal condition is just over or equal to 50%. Comparing the value of T, the monitoring capability of different control charts can be analyzed and compared.

3.3. Product sales tracking time and monitoring time

The production and sales of products are continuous to the manufacturer, and most of them have a time delay, which could be several months or even up to years, from production to sales. For example, the average sales delay of automobile is about 25 weeks, which implies that the automobiles will be completely sold out in about 25 weeks after their production. For the monitoring model proposed by this study, it is suggested that products of different production and sales time monitored as a whole index at different monitoring stages. If the reli-

ability of products starts to exhibit abnormalities after some certain production stages, owing to sales delay, it is not obvious in the early monitoring stages because of the small proportion of these products, as shown in Fig 1. Then, the initial reliability variance will difficult to be detected as soon as possible. To deal with this problem, a reasonable limit of product sales tracking time in the proposed model is necessary in this study.

In addition, the products may have the same claim rate at different stages during the product warranty, then the theoretical claim quantity at different monitoring stages is only dependent on the quantity of products monitored, which is quite simple and will not be expanded in this study. More commonly, for product such as automobile, its claim rates at different stages are not the same, which is often fitted with Weibull distribution or lognormal distribution. For this circumstance, a long monitoring time of the products will result in poor detection power of abnormal conditions for the newly listed products, which is caused by the relatively small proportion of them under monitoring. Therefore, it is also necessary to limit the monitoring time for the products according to its claim rate function.

Suppose A and B to be the limit of product sales tracking time and monitoring time respectively. Taking monitoring stage l as an example,the actual claim quantity and expected claim quantity will be adjusted to:

$$o_l = \sum_{\substack{i+j+k-2=l \\ j\le A, k\le B}} R_{ijk}$$

$$e_l = \sum_{\substack{i+j+k-2=l, \\ j\le A, k\le B}} N_{ij} * \gamma_k^0$$

$$\forall i+(j-1)+(k-1)\le M \tag{11}$$

A limit of both product sales tracking time and monitoring time willnot only improve the early-warning ability of the proposed model, but also facilitate the process of data collection and modeling effectively.

4. Case study

In order to illustrate and test our proposed monitoring model for products sold with two-dimensional warranty, a case study is carried out in this section by using the warranty data of an automobile manufacturing enterprise in China. Firstly, the main parameters of the proposed model are set or estimated from the enterprise in subsection 4.1. Secondly, through analyzing the response time of the proposed model, the validity and applicability of the proposed monitoring model have been demonstrated. Thirdly, comparing the response time for abnormal conditions, it has been verified that a reasonable combination of A and B can effectively improve the early-warning ability of the proposed control chart. Fourthly, comparing with CUSUM control chart, the monitoring capability of the proposed model has been proven in subsection 4.4. Finally, an actual failure mode has been monitored by both CUSUM control chart and the proposed control chart in this study.

4.1. Parameter estimation

Parameter estimation for Accelerated failure time (AFT) and probability density function (PDF) of product usage rate have been elaborately studied in literature [7, 19], which isn't described in detail in this study. Considering one week as the smallest unit of time for the independent stage, the prespecified number of stages for monitoring is M=40, product sales tracking time is $A=20$, and monitoring time is $B=20$ respectively. The overall false alarm rate up to stage M is 0.05, and the sales of automobiles within 20 weeks is uniform distribution over $[4500,5500]$.Besides, the usage rate r is also uniform distribution over $[0.5,3.5]$. For the nominal usage rate $r_0=2$, the time to the first failure of product is Weibull distribution with dimension parameter $\alpha_0>0$ and shape parameter $\beta>0$, then the cumulative distribution function (CDF)of the time to the first failure is:

$$F_d(x:\alpha_0,\beta)=1-e^{-(x/\alpha_0)^{\beta}} \tag{12}$$

With the accelerating factor $\gamma=1.2$, which is calculated through AFT. The hazard function of the time to the first failure for product with nominal usage rate r_0 and actual usage rate r are calculated respectively, which are:

$$z_D(x)=\beta(\frac{1}{\alpha_0})^{\beta}t^{\beta-1}$$

$$z_r(x)=\left(\frac{r}{r_0}\right)^{\gamma\beta} z_D(x) \tag{13}$$

4.2. Model applicability analysis

For Weibull distribution with shape parameter β, when $\beta>1$, the hazard rate function is increasing with time, while when $0<\beta>1$, the hazard rate function is decreasing [18]. In other words, when $\beta\neq1$, the product claims data in the whole warranty obey non-homogeneous Poisson distribution (NHPP). In order to test the monitoring capability of the proposed model for different hazard rate function, two different hazard functions, which are α_0=200, β=2, and α_0=12500, β=0.5, are chosen to analyze the response time of the proposed control chart. For the first instance, the allocation of false alarm probability and the

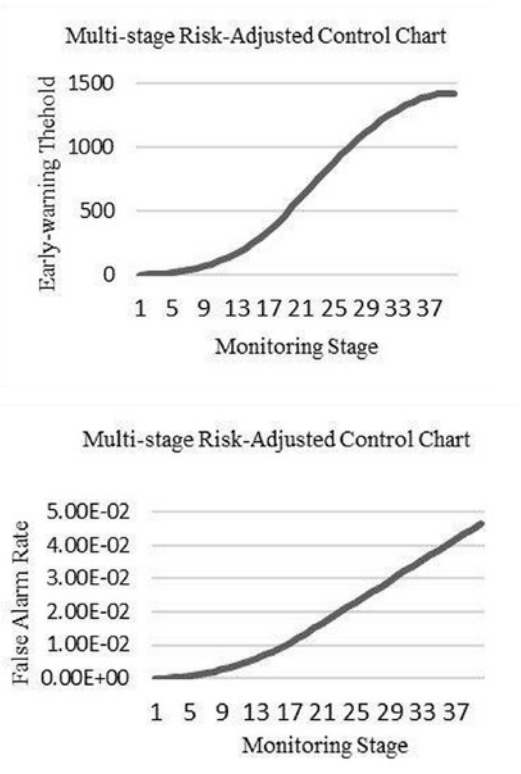


Fig. 3. Early-warning threshold and false alarm rate of each monitoring phase

Table. 1. Early-warning ability of two different distributions

stage	α_0=200, β=2			α_0=12500, β=0.5		
	1.2	1.3	1.4	1.2	1.3	1.4
10	0.003232	0.003236	0.003241	0.005497	0.006575	0.008152
12	0.004999	0.00511	0.005227	0.030431	0.07723	0.173059
14	0.008096	0.008862	0.00976	0.179185	**0.515428**	**0.852389**
16	0.01332	0.016624	0.021111	**0.587684**	0.967233	0.99982
18	0.0249	0.038204	0.059221	0.939524	0.99997	1
20	0.048138	0.091951	0.168141	0.998978	1	1
22	0.099147	0.227474	0.436915	0.999999	1	1
24	0.20822	**0.503929**	**0.81835**	1	1	1
26	0.420305	0.844438	0.990402	1	1	1
28	**0.720033**	0.989995	0.999988	1	1	1
30	0.944992	0.999972	1	1	1	1
32	0.997092	1	1	1	1	1
34	0.999975	1	1	1	1	1
36	1	1	1	1	1	1
38	1	1	1	1	1	1
40	1	1	1	1	1	1

threshold for different monitoring stages are calculated with $z_r(x)$ and α_l by eq. (7) and eq. (8), as shown in Fig. 3.

Assuming that the warranty claim rates of products produced after 10 weeks has become 1.2-1.4 times of the original standard. The response time for these abnormal conditions of the proposed model is analyzed in table 1

As shown in table 1, the theoretical alarm probabilities for abnormal conditions at monitoring stages from 10 to 40 have been calculated. It is obvious that the distribution of warranty claim rate has an important influence on the response time of the proposed model, which should to be considered in the design of the proposed model. Through timely and effectively detecting the claim rate change occurring during monitoring stages, the experimental results show that the proposed control chart can be applied to the monitoring and early-warning of product claim rate with different distribution function.

4.3. Parameter sensitivity analysis

In order to test the influence of product sales tracking time A and monitoring time B on the monitoring capability of the proposed model, various parameter combinations of A and have been considered for monitoring the first example mentioned above, the monitoring results are shown in table 2.

As can be seen in table 2, for abnormal conditions occurred after 10^{th} stage, the alarm probabilities of the 10^{th}–30^{th} monitoring stages have been calculated. The experimental results show that different combinations of A and B can effectively influence the alarmability of the proposed model. Firstly, the small value of can reduce the risk of a weak early-warning ability caused by sales delay,but it also need sufficient data to improve the confidence level of the proposed model. Therefore, we set the time when sales percentages of products up to 80% as the value ofin this study, which not only ensures that there is enough warranty data, but also greatly reduces the deficiency of the weak warning ability caused by sales delay. Secondly, the value of has the same influence on the early-warning ability of the proposed model, the choice of which depends on the hazard rate function of the product during the warranty period. When the product warranty claim rates decrease throughout the warranty period, it can effectively improve the early-warning ability to shorten the monitoring time. On the contrary, when the product claim rates increase throughout the warranty period, reasonably prolonging the monitoring time is necessary to ensure that the warranty claims data are large enough for monitoring. Therefore, the value of should be adjusted according to the hazard rate function of the time to the first failure of the product. On the whole, a reasonable combination of product sales tracking time and monitoring timecan effectively reduce the risk of weak warning ability caused by sales delay or fluctuating claim rates.

Table 2. Results of Early-warning Ability under different A and B

Monitoring Stage	α_0=200, β=2				α_0=12500, β=0.5			
	B=20		B=10		B=20		B=10	
	A=10	A=20	A=10	A=20	A=10	A=20	A=10	A=20
10	0.005896	0.003232	0.005896	0.003232	0.008942	0.005497	0.008942	0.005415
12	0.008655	0.004999	0.008986	0.004896	0.043345	0.030431	0.043449	0.029906
14	0.012558	0.008096	0.012866	0.0079	0.249646	0.179185	0.261572	0.187031
16	0.018755	0.01332	0.020244	0.014252	**0.745238**	0.587684	0.786577	0.618208
18	0.030785	0.0249	0.039151	0.030662	0.990549	0.939524	0.996431	0.958025
20	0.055415	0.048138	0.101427	0.074181	0.999992	0.998978	1	0.999678
22	0.11061	0.099147	0.248367	0.166329	1	0.999999	1	1
24	0.23438	0.20822	0.477935	0.312786	1	1	1	1
26	0.475417	0.420305	**0.70817**	**0.505569**	1	1	1	1
28	**0.781122**	**0.720033**	0.860335	0.709367	1	1	1	1
30	0.967884	0.944992	0.934663	0.878262	1	1	1	1
32	0.99892	0.997092	0.969434	0.966794	1	1	1	1
34	0.999992	0.999975	0.985701	0.994366	1	1	1	1
36	1	1	0.993311	0.999376	1	1	1	1
38	1	1	0.996871	0.999947	1	1	1	1
40	1	1	0.998536	0.999996	1	1	1	1

Table 3. Results of Early-warning ability under different Control Charts

Monitoring Stage	Multi-stage Risk Adjustment Control Chart			CUSUM Control Chart		
	1.2	1.3	1.4	1.2	1.3	1.4
10	0.003232	0.003236	0.003241	0.046123	0.046151	0.04618
12	0.004999	0.00511	0.005227	0.051691	0.051992	0.052308
14	0.008096	0.008862	0.00976	0.054124	0.054926	0.05584
16	0.01332	0.016624	0.021111	0.055111	0.056457	0.058189
18	0.0249	0.038204	0.059221	0.055499	0.057335	0.060088
20	0.048138	0.091951	0.168141	0.055636	0.057853	0.061846
22	0.099147	0.227474	0.436915	0.0557	0.058322	0.064544
24	0.20822	**0.503929**	0.81835	0.055747	0.059054	0.071637
26	0.420305	0.844438	0.990402	0.055812	0.061231	0.10218
28	0.720033	0.989995	0.999988	0.055984	0.071743	0.251154
30	0.944992	0.999972	1	0.056638	0.126209	**0.671404**
32	0.997092	1	1	0.058899	0.303362	0.969075
34	0.999975	1	1	0.0665	**0.637179**	0.99981
36	1	1	1	0.088608	0.916635	1
38	1	1	1	0.147474	0.995197	1
40	1	1	1	0.280899	0.999966	1

4.4. Early-Warning Ability Analysis

Traditional CUSUM control chart has been used in the monitoring and early warning of products sold with one-dimensional warranty, which can also be adjusted to monitor products sold with two-dimensional warranty. In order to test the early-warning ability of the proposed model in this study, a comparison of response timefor the above-mentioned control charts has been carried out in this subsection.

As seen in table 3, the first instance in subsection 4.2, has been chosen as an example in this analysis. The hazard rate function of the time to the first failure is Weibull distribution with parameters, and it is known that the warranty claim rate of product produced after 10 weeks has become 1.2-1.4 times of the original standard. On the condition that the overall false alarm rate up to 40 weeks is 0.05, the alarm probabilities of the above two control charts at different monitoring stages have been calculated. It is obvious to find that comparing to traditional CUSUM control chart, the alarm probabilities for abnormal conditions of the proposed model are greater at the following monitoring stages, which have reached to 50% as soon as possible. The result means that the response time of the proposed model for abnormal conditions is shorter than traditional CUSUM control chart. However, we have only analyzed the response time for relatively small change of warranty claims data in table 3, extensive experimental results show both traditional CUSUM control chart and the proposed control chart in this study have a shorter response time for huge change of claim rates, which isn't discussed in detail in this study.

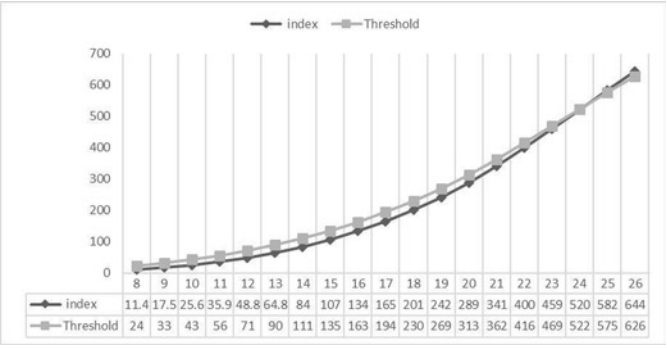


Fig . 4. Multi-stage risk-adjusted control chart

4.5. Example analysis

In this subsection, the proposed control chart and CUSUM control chart are both used to monitor the actual claims data for a common failure mode in 2014.For the normal usage rate , the hazard rate function of the time to the first failure is fitted by Weibull distribution with ,and the accelerating factor of AFT is $\gamma = 1.2$, The prespecified monitoring stages, the product sales tracking time and all other parameters of the proposed model are the same as studied in subsection 4.1.It is known that the claim rate of products has been increased about 1.2 times of the original standard from week 8, which is caused by the change of raw material. The real monitoring performance of the proposed model and CUSUM control chart are shown in Fig 4 and Fig 5.

As seen in Fig. 4 and Fig. 5, both the proposed control chart and CUSUM control chart have been applied for monitoring actual warranty claims data. It is obvious to find that the monitoring thresholds of the proposed control chart are changing stage by stage, while the monitoring thresholds of CUSUM control chart are constant among all the monitoring stages.

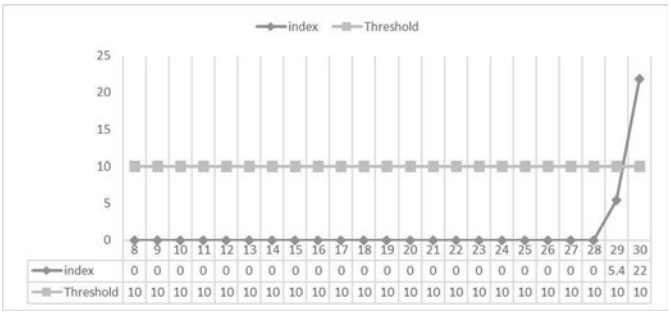


Fig . 5. Traditional CUSUM control chart

For the abnormal conditions of claim rates occurring from week 8, the proposed control chart needs 16 weeks to detect the change and raise the alarm, while the traditional CUSUM control chart needs23 weeks. The multi-stage risk-adjusted control chart proposed by this study, which has a relatively shorter response time for the abnormal change of claim rate, can effectively apply for the monitoring and early-warning of two-dimensional warranty claims data.

5. Summary

Early detection of reliability problems can aid manufacturers to reduce the associated warranty costs and improve customer satisfaction and brand image. For products sold with two-dimensional warranty, a multi-stage risk-adjusted control chart is provided by this study, which can detect and warn the quality or reliability issues of products as soon as possible. However, in consideration of the fact that the maintenance strategies have a significance influence on the claim rate distribution during the warranty period, only minimal repair strategy, for which the monitoring of different stages is independent from each other, has been considered in this study. It is, therefore, a great challenge to design a control chart for the monitoring of time-related claims data, which will be faced by the application of other maintenance strategies.

Acknowledgments

The authors would like to thank the work of the reviewers for their contribution to the quality of this paper. Part of this research is funded by the National Natural Science Foundation of China (No. 71532008)

References

1. Altun M, Comert S V. A change-point based reliability prediction model using field return data. Reliability Engineering & System Safety 2016; 156: 175-184, https://doi.org/10.1016/j.ress.2016.07.024.
2. Asadzadeh S, Aghaie A, Shahriari H. Using frailty models to account for heterogeneity in multistage manufacturing and service processes. Quality & Quantity 2014; 48(2): 593-604, https://doi.org/10.1007/s11135-012-9789-x.
3. Baik J.: Reliability assessment based on two-dimensional warranty data and an accelerated failure time model. International Journal of Reliability & Safety 2008; 2(3): 190-208, https://doi.org/10.1504/IJRS.2008.021064.
4. Grabert M, Prechtel M, Hrycej T, et al. An Early Warning System for Vehicle Related Quality Data. In: Advances in Data Mining, Applications in Image Mining, Medicine and Biotechnology, Management and Environmental Control, and Telecommunications, Industrial Conference on Data Mining, 2004: 88-95, https://doi.org/10.1007/978-3-540-30185-1_10.
5. Haq A, Al-Omari A I. A new Shewhart control chart for monitoring process mean based on partially ordered judgment subset sampling. Quality & Quantity 2015; 49(3): 1-18, https://doi.org/10.1007/s11135-014-0042-7.
6. Honari B, Donovan J. Early Detection of Reliability Changes for a Non-Poisson Life Model Using Field Failure Data.In: Reliability and Maintainability Symposium 2007: 346-349, https://doi.org/10.1109/RAMS.2007.328137.
7. Iskandar B P, Murthy D N P. Repair-replace strategies for two-dimensional warranty policies. Mathematical & Computer Modelling 2003; 38(11-13): 1233-1241,
8. J. F. Lawless, M. J. Crowder, K.- A. Lee. Monitoring Warranty Claims With Cusums. Technometrics 2012; 54(3): 269-278, https://doi.org/10.1080/00401706.2012.680373.
9. Lee S H, Lee S J, Moon K I. Application of Fuzzy Feedback Control for Warranty Claim. International Journal on Information 2011; 351(11): 279-288, https://doi.org/10.1007/978-3-642-19953-0_28.
10. Lee S H, Seo S C, Yeom S J, et al. A Study on Warning/Detection Degree of Warranty Claims Data Using Neural Network Learning. In: International Conference on Advanced Language Processing and Web Information Technology. IEEE 2007: 492-497, https://doi.org/10.1109/ALPIT.2007.82.
11. Na C S, Jung B S. AHP-Based Determination of Warning Grade in a Warranty Claims 2010; 11 (12): 5097-5106.
12. Odanaka T. Control chart and stochastic control processes. Computers & Mathematics with Applications 1991; 21(11-12): 65-78, https://doi.org/10.1016/0898-1221(91)90108-G.
13. Shahanaghi K, Noorossana R, Jalali-Naini S G, et al. Failure modeling and optimizing preventive maintenance strategy during two-dimensional extended warranty contracts. Engineering Failure Analysis 2013; 28(3): 90-102, https://doi.org/10.1016/j.engfailanal.2012.09.006.
14. Varnosafaderani S, Chukova S.: A two-dimensional warranty servicing strategy based on reduction in product failure intensity. Computers & Mathematics with Applications 2012; 63(1), 201-213. https://doi.org/10.1016/j.camwa.2011.11.011
15. Veiga P, Mendes L, Lourenço L. A retrospective view of statistical quality control research and identification of emerging trends: a bibliometric analysis. Quality & Quantity 2016; 50(2): 1-20, https://doi.org/10.1007/s11135-015-0170-8.
16. Wu H, Meeker W Q. Early Detection of Reliability Problems Using Information from Warranty Databases. Technometrics 2002; 44(2): 120-133, https://doi.org/10.1198/004017002317375073.
17. Wu S. Warranty Data Analysis: A Review. Quality & Reliability Engineering International 2012; 28(8): 795-805, https://doi.org/10.1002/qre.1282.
18. Yeh R H, Chen C K. Periodical Preventive-Maintenance Contract for a Leased Facility with Weibull Life-Time. Quality & Quantity 2006; 40(2): 303-313, https://doi.org/10.1007/s11135-005-8095-2.
19. Zhou C, Chinnam R B, Korostelev A. Hazard rate models for early detection of reliability problems using information from warranty databases and upstream supply chain. International Journal of Production Economics 2012; 139(1): 180-195, https://doi.org/10.1016/j.ijpe.2012.04.007.
20. ZH Cheng, Yang Z Y, Zhao Z, et al. Preventive maintenance strategy optimizing model under two-dimensional warranty policy. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17(3): 365-372, https://doi.org/10.17531/ein.2015.3.6.

Fangqi DONG
Zixian LIU
Yujie WU
Jianhong HAO
College of Management and Economics
Tianjin University
No.92, Weijin Road, Nankai District, Tianjin, China

E-mails: dfq0330@126.com, liuzixian@tju.edu.cn, wuyujie@tju.edu.cn, haojianhong@catarc.ac.cn

Article citation info:
KILIAN-BŁAŻEJEWSKA E. The influence of selected factors on the failure rate of the water pipelines located on the area of the impact of mining tremors. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 308–317, http://dx.doi.org/10.17531/ein.2018.2.18.

Ewelina KILIAN-BŁAŻEJEWSKA

THE INFLUENCE OF SELECTED FACTORS ON THE FAILURE RATE OF THE WATER PIPELINES LOCATED ON THE AREA OF THE IMPACT OF MINING TREMORS

WPŁYW WYBRANYCH CZYNNIKÓW NA INTENSYWNOŚĆ USZKODZEŃ PRZEWODÓW PODSYSTEMU DYSTRYBUCJI WODY, ZLOKALIZOWANYCH NA TERENIE ODDZIAŁYWANIA WSTRZĄSÓW POCHODZENIA GÓRNICZEGO*

The article presents the influence of the selected factors, including mining tremors, described by the parameter PGV_{Hmax} on the failure rate of the water distribution pipelines. In created multiple regression models, the following independent variables were used: diameter and material from which the pipes were made, operation time without failure, the age of pipes, the value of pressure and PGV_{Hmax}. The values of PGV_{Hmax} in places with damaged water pipelines were determined by constructed the seismic wave propagation models. The analysis was carried out for a random sample of all water pipelines and extracted from this sample new groups: steel and gray cast iron pipes, their diameters, diameters and materials, and their construction time.

***Keywords**: water distribution system, mining tremors, failure rate of water pipelines.*

W artykule przedstawiono badania wpływu wybranych czynników, w tym wstrząsów pochodzenia górniczego, opisanych za pomocą parametru PGV_{Hmax} na intensywność uszkodzeń przewodów podsystemu dystrybucji wody. Jako zmienne niezależne w utworzonych modelach regresji wielorakiej przyjęto: średnicę i materiał z którego wykonane są przewody, czas pracy bezuszkodzeniowej, wiek przewodów, wysokość ciśnienia i PGV_{Hmax}. Wartości PGV_{Hmax} w miejscach występowania awarii przewodów wodociągowych, wyznaczone zostały na podstawie zbudowanych modeli propagacji fali drgań w ośrodku gruntowym. Analiza przeprowadzona została dla próby losowej obejmującej sumarycznie wszystkie przewody sieci wodociągowej magistralnej, rozdzielczej i przyłącza oraz dla wyodrębnionych z tej grupy prób losowych obejmujących: przewody zbudowane ze stali i z żeliwa szarego, przewody z uwzględnieniem ich średnicy, średnicy i materiału oraz z uwzględnieniem czasu ich budowy.

***Słowa kluczowe**: podsystem dystrybucji wody, wstrząsy górnicze, intensywność uszkodzeń przewodów sieci wodociągowej.*

1. Introduction

Water supply systems are a key part of the technical infrastructure for the functioning of the economy and society. The main tasks of water supply systems are: supply of water in the required quantity, quality and under appropriate pressure [51]. Maintaining the functionality and continuity of operation of these systems require ensuring their appropriate level of safety [30,53,34,46], which is obligated, among others, by Council Directive 2008/114/EC [8] and the Act of 26 April 2007 on crisis management [50]. Critical infrastructure protection procedures, to which water supply has been qualified, as well as procedures for determining the reliability and security of water supply to recipients, include in the first place identification of all hazardous factors [7,9,4,34,35,43,44,45] .

Many analyzes of the impact of various factors on the damage intensity of the water distribution system have described in the literature [5,15,17,29,35,40]. The following factors have analyzed: diameter, material, temperature of water in the water pipelines, temperature of ambient, length of pipes, average rainfall, type of pipes, depth of pipes, operation time without failure, pressure, type of ground, aggressiveness of the water and soil as well as the impact of mining, defined only as a qualitative variable (zero - one). The influence of mining tremors on the water pipe failure has not been analyzed yet.

Accordingly, in order to complete the analysis of pipe failures and reliability of systems located in mining areas, it is necessary to include in the analyzes a accordingly described factor, informing about the impact of mining tremors.

Mining exploitation is very often carried out in Poland under highly urbanized areas, densely armed with underground, linear technical infrastructure, like networks of: water, sewage, gas, heating, telecommunication. This creates numerous threats both for surface cubage objects and for the underground utilities network.

Water supply networks are characterized by various constructional and material solutions, connection methods, their age and technical condition are different. For these reasons, the pipelines are characterized by a varied dynamic response to additional loads resulting from the mining deformed area and surface vibrations induced by mining tremors. In addition, it is very difficult to identify the technical condition of individual pipelines, like their real load resistance, which makes it very difficult to carry out advanced analyzes.

For seismic areas, dependencies describing the impact of earthquakes with specific parameters on the damage intensity of water supply pipelines have been developed. The conducted research [23.39] has shown that PGV (Peak Ground Velocity) better correlates with pipelines damage caused by earthquakes than PGA (Peak Ground Acceleration). It was shown that PGV is directionally correlated with maximum ground deformations - the main cause of pipelines damage during propagation of the seismic wave. PGA better correlates with inertia forces that do not affect underground objects. Pineda - Poras and Ordaz [42] showed that the model of prediction of failure intensi-

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

ty of water network pipes, based on the PGV parameter, overestimates the number of failures caused by earthquakes magnitude 8.0 - 8.1 and smaller and underestimate the number of pipe failures for stronger earthquakes. The parameter defined like PGV^2/PGA describes the relationship between earthquakes and the failure intensity of pipelines in a better way.

In addition to building of empirical models, many other studies have been carried out on the impact of earthquake parameters on pipelines buried in the ground. For example, Wang and Cheng [52] demonstrated by constructing a static numerical model that the behavior of the earthquake - laden pipeline depends on the wave transition time as well as the inhomogeneity of the ground in which the pipe is buried. Takada and Tanabe [48] have developed a three - dimensional static numerical model for pipelines and their connections load by large earthquakes. It was shown that the propagation of waves vibration at individual locations is characterized by peak values as well as by the speed of wave propagation. O'Rourke and Liu [37] analyzed the deformations and the curvature of the ground formed during the earthquake wave propagation, depending on the various ground conditions, on the formation of surface deformations. Temporary surface deformations of the ground are considered the most important from the point of view of the impact of seismic waves propagation on underground linear objects, like pipelines and tunnels. Lee [33] analyzed the history of earthquakes and on its basis investigated the impact of earthquake parameters on a buried gas pipeline in the ground including: type of pipeline material (plastic, brittle), diameter, ground properties, depth of gaz pipeline including location (residential and industrial area, public road, port areas - pipeline submerged, etc.). O'Rourke and Ayala, [36] have shown that pipelines made of flexible materials are more resistant to earthquakes than fragile pipelines. Similarly, Bouziou and O'Rourke [4] pointed out that the most resistant material of water pipelines to earthquake that occurred on 22 February 2011 in Christchurch was polyvinyl chloride (PVC). O'Rourke and others [38] analyzed the behavior of the ground at the interface with the pipeline under the influence of ground deformation caused by earthquakes. The tests have shown that it is necessary to change the design assumptions (static and strength calculations) for the forces occurring from the bottom of the pipeline.

The impact of earthquakes on underground, linear objects of technical infrastructure is relatively well described in many literature items around the world. The problem of the models of seismic waves propagation is still the accuracy of models, which are affected by a large variety of factors contributing to the creation of a specific impact of earthquakes on the surface and the objects.

Considering the significant differences between earthquakes and mining tremors, the implementation of dependencies developed for the earthquakes may be difficult and requires accordingly analysis. However, this does not change the fact that rich studies on seismic waves can provide an excellent source of knowledge for analysis of the impact of mining tremors on underground, linear technical objects.

Differences between earthquakes and mining tremors are caused among other by [55]:

1. size and intensity - strong mining tremors can only be compared with small earthquakes,
2. duration of kinematic extort - the duration of mining tremors is a few seconds, whereas an earthquake can last 10 - 30 seconds,
3. properties of spectral records - which is the reason for significant differences in the impact of phenomena on building objects,
4. peaks of ground movement (PGA, PGV, PGD - permanent deformation of the ground),
5. differences between the vertical and horizontal components of ground vibrations.

It is relatively easy to identify damage of pipelines caused by earthquakes. This is possible due to the scale of the phenomenon's impact, as well as the ability to correlate its occurrence with the flow increase (based on data from monitoring system), informing about the increase of the number of water distribution system failures. In addition, the location of the failure can be simpler, due to the size and intensity of the water outflow from the damaged pipes.

In the case of pipeline failures that may be caused by the impact of mining tremors, the situation is more complicated because it can only refer to unit damages, which due to insufficient research, it is difficult to distinguish from a group of other causes of failures.

In the literature, it has often pointed out that the impact of mining tremors on linear objects buried in the ground is not sufficiently known [20, 21, 22] or is treated as secondary [28]. The paper [22] indicates that unambiguous determination of the impact of mining tremors on pipeline damage was difficult, due to the possible time shift of the cause and effect of damage. In spite of this, publications presenting research results (mainly numerical or static - endurance analyzes) concerning the impact of mining tremors on linear objects are met.

Dulińska [11] examined the effect of surface vibrations on the structure of the gas pipeline, by calculating the dynamic response of the object from strong tremors. The analyzes considered a 100 m section of the gas pipeline, three variants of the soil were considered: dry sands, clay loons and sandy gravels. Elastic cooperation of the pipe with soil in vertical and horizontal direction was assumed. The model of uneven kinematic extortion was tested (at each point of the pipeline at the same moment there is different kinematic exclusion). The analysis showed that higher dynamic responses to kinematic excitations are obtained for lower wave propagation velocities. The frequency structure of ground vibrations affects the size of the gas pipeline response, even at approximated maximum vibration amplitudes at different frequency structures. Max stresses in the gas pipeline under the influence of the tremor were 18 MPa or 10% steel strength. In the works [24] and [25], attention was paid to the necessity of each time analyzing the impact of ground vibrations on gas pipelines and to examine their dynamic response to the existing vibrations. This is to prevent possible consequences of damage to old steel pipelines. Dulińska and Jasińska [12] investigated using numerical methods the maintenance of a 100 m steel pipeline on concrete supports during a seismic tremor. The propagation of the wave vibration along the pipeline was analyzed. In the pipeline, depending on the assumed model, cracks occurred at the location of the concrete supports. Boron P., Dulińska J., [3] analyzed the numerical response of joints (including bolts, flanges and retaining block) of steel pipelines to the impact of seismic tremor. Obtained deformation of the bolts indicated the occurrence of high values of bending forces at the pipeline connection points. The paper [25] presents an analysis of impact of tremor with energy $8 \cdot 10^7$ J on steel gas pipe DN200, low pressure, equipped with gland compensators and characterized by good technical condition. As a part of the work, a static - strength and kinematic analysis was carried out for a section of a straight gas pipeline with a length of 50 m and a 90° arc. The conducted analysis showed that the forces (displacements, tensions and strains) reacting on the gas pipeline take values lower than those caused by continuous deformation of the ground. Kurzeja J. [26] investigated the impact of mining tremors on the Piekary Śląskie (Poland) A1 motorway junction. The analysis indicated that there is no negative impact of mining tremors on the motorway. The inference was based on systematic observations of the state of the motorway.

The aimed of the research presented in this article is the identification of the relationship between mining tremors described by the PGV_{Hmax} and the damage intensity of water distribution subsystem, taking into account: diameter and material from which pipelines were made, construction time, operation time without failure and water

pressure. These are the first results of this type of analysis published in the literature.

The developed dependencies can be used to forecast damage to water supply network pipes in the analyzed region, under the influence of dynamic loads of mining operations.

2. Methodology of the research

The subject of the research was the water distribution subsystem, located in the area with the impact of dynamic stresses, caused by mining tremors. The research included analysis of all failures of the water distribution subsystem and recorded mining tremors with energy $E \geq 10^6$ J, occurring in 2011-2014.

The scope of research included:

1. determination of the local attenuation relations for mining tremors described by the PGV_{Hmax} parameter,
2. determining the value of PGV_{Hmax} in places where the failures of water pipes occoured,
3. construction of multiple regression models to identify factors affecting the damage intensity of the water distribution subsystem.

2.1. Local attenuation relation for the mining tremors describes by the PGV_{Hmax}

In literature [1, 6, 14, 19, 26], many attempts have been made to determine the empirical relationship between seismic parameter and factors affecting it's size. These dependencies were determined both for mining tremors and for earthquakes.

In 1999, Si H. and Midorikawa S. [48] developed a regression model in which the following factors were adopted as a factors influencing the magnitude of the seismic effect: energy of tremor, the mechanism of the source of tremor, the distance from the source of the tremor, and for the first time: the geological structure of the soil.

The model has the following form:

$$logA = b - logR - k \cdot R \tag{1}$$

where:

A – seismic parameter for example: PGV (Peak Ground Velocity), PGA (Peak Ground Acceleration) or PGD (Peak Ground Displacement),

R – the distance from the source of the tremor, [m],

$$R = \sqrt{d^2 + h^2} \tag{2}$$

d – distance from the hypocenter, [m],

h – pseudo depth determined by regression which provides non - linearity of the relationship for a small distance from the source, [m],

k – coefficient that accepts the following values: for PGV – 0,002 and for PGA – 0,003, [-],

b – offset factor, [-],

$$b = a \cdot \log E + c \cdot R + \sum d_i \cdot S_i + e + \varepsilon \tag{3}$$

where:

a, c, d, e – regression parameters estimated from seismometric measurements, [-],

E – energy of tremor, [J],

Si – ground type, quality variable, distance di dependence on the number of ground types i, adopted in this research in accordance with Eurocode 8 (Tab. 1) [13], [-]

ε – random factor, [-].

R – like in equation (1)

The parameter adopted in this analysis for the description of ground vibrations (seismic parameter) was the maximum amplitude of horizontal vibrations velocity, determined as the horizontal maximum of the vibration length of the ground vibration (PGV_{Hmax}).

Model (1) became the basis for the development of the attenuation relations for the area in which the water distribution subsystem is located and for determining the seismic parameter value in places where the water supply failures occured.

Table 1. Ground type classification in accordance with Eurocode 8 [13]. $V_{S,30}$ - average speed of propagation of S waves in soil layers up to a depth of 30 m.

Ground type	Description of stratigraphic profile	$V_{s,30}$, m/s
A	Rock or other rock-like geological formation, including at most 5m of weaker material at the surface	>800
B	Deposits of very dense sand, gravel or very stiff clay, at least several tens of meters in thickness, characterized by a gradual increase of mechanical properties with depth	360-800
C	Deep deposits of dense or medium-dense sand, gravel or stiff clay with thickness from several tens to many hundreds of meters	180-360
D	Deposits of loose-to-medium cohesion less soil (with or without some soft cohesive layers) or of predominantly soft-to-firm cohesive soil	<180
E	A soil profile consisting of a surface alluvium layer with Vs values of type C or D and thickness varying between about 5m and 20m, underlain by stiffer material with Vs>800 m/s	-

2.2. Determination of PGV_{Hmax} values in places where water supply was failure

The procedure required assigning to each mining tremors the water supply failures caused by them. For this purpose, it was necessary to adopt the following assumption: mining tremor may be the cause of any water supply failure occurring up to 30 days after the tremor.

The necessity of adopting this assumption was caused by the type of the date of water supply failures owned by the water supply company. The analyzed database had only information about time of start repair the pipelines after failure. For the purpose of the study, it was important to obtain information about the time of failure. The moment of failure due to the large number of unknowns and the low accuracy of measuring devices is not possible to determine precisely at the moment. According to the above, the 30 days established is a kind of buffer in which the failure could have started.

In the research, no spatial restriction concerning the maximum distance between the tremor and the failure of the water pipe was accepted, due to the unknown effective range of impact of mining tremors on underground linear objects.

Determination of the PGV_{Hmax} value in the place where the water pipe failures, using the model created on the basis of equation (1), required: conducting ground classification in the place of individual failures, according to Eurocode 8 [13] and determining the epicentric

distance (between mining tremors and water supply failures) based on geographical coordinates.

In the case when the same water supply failure was assigned to the various mining tremors, the "tremor-failure" pair was left in the analysis in which the PGV_{Hmax} parameter assumed a higher value. Finally, 993 "tremor - failure" pairs were obtained. Fig. 1 shows the spatial distribution of all analyzed mining tremors and failures of the water distribution subsystem.

Fig. 1. The spatial distribution of mining tremors (of $E > 10^6$J) and failure of the water supply network in 2011-2014

2.3. The influence of selected factors on the damage intensity of the water distribution subsystem

In the next stage of the research for individual water pipes, characterized by parameters: number of failures, diameter, material, age, etc., the average value of PGV_{Hmax} was assigned from all water supply failure locations on the same pipeline. In the research, through the water pipe were understood, the sections of the main and distribution water supply network and connections, characterized by: diameter, material and time of construction. The analysis of the impact of selected factors on the failure intensity of water distribution subsystems was conducted for a group of 792 pipes.

In the multiple regression model, the impact of selected factors on the unit failure intensity of water distribution subsystem, the following variables were adopted:

- unit failure intensity rate (λ) - dependent variable of the built regression models,
- and independent variables:
- material (M) and diameter of individual pipelines (DN),
- average pressure in the place of failure on a pipeline (P),
- average age of the pipeline at the time of damage (W),
- operation time without failure of the pipeline in the analyzed time (Tp),
- PGV_{Hmax}.

The most commonly parameters used for assessing the failure frequency of water pipes are: stream of failure ω (t) or damage intensity λ (t). Assuming that the stream of failures is a stream: without consequences (damages occurring at particular times Δt are independent), single (probability of more than one damage on any small part of time goes to zero) and stationary (probability of k damage in the time interval Δt, depends only on this interval, and does not depend on the position on the time axis), the equality of this parameter with the intensity of damage λ (t) = ω (t) = const is obtained [50.52]. Based on the presented assumption, in the conducted research the failure frequency of the water pipes was described using the failure intensity indicator.

The unit failure intensity λ(Δt) for linear elements (pipelines) is expressed in equation [31.53]:

$$\lambda(\Delta t) = \frac{n(\Delta t)}{L\,\Delta t} \quad [\text{damage}/(\text{km}\cdot\text{a})] \tag{4}$$

where:

- λ (Δt) – unit failure intensity to linear objects, [damaged/(km a)],
- n (Δt) – number of failures in the time interval Δt,
- L – average length of analyzed pipelines in the time interval Δt, [km],
- Δt – the time interval for which the observation period has divided, [a].

Failure intensity of the analyzed water supply network pipelines was λmaxs = 200 damaged/(km·4 years), median = 8.8 damaged/(km·4 years), and the failure intensity of connection λmaxp = 2000 damage/(km·4 year), median = 93.9 damage/(km·4 years).

The water supply system can be treated as a renewable (cyclically occurring states of work and disability), from the point of view of reliability theory. It is called the „binary model with non-zero renovation" [31.53]. In the case of non-damaged operation time has an exponential distribution:

$$T_p = \frac{1}{\lambda} \tag{5}$$

where:

- T_p – average operation time without failure of the linear object, [years]
- λ – damage intensity of the linear object, [damage / (km year)],

In the analysis, the operation time without failure was within the range (26 - 730) days. It should be noted that most of the pipelines (in particular the connections) were damaged once in the analyzed time and the operation time without failure assumed the same value equal 730 days (median).

The pressure at the place where the water pipes failure was calculated based on the knowledge (from the monitoring data) of the pressure at the point supplying the pipeline or the area of the water supply network where the pipeline was located and the height difference between the supply point and the place of failure (read from numeric terrain model). For individual pipelines, the pressure was determined as the arithmetic average of the pressure from all failures locations on a same pipeline.

The pressure for all analyzed pipelines was within the range (13.39 - 86.87) m, while the median was: 44.48 m.

Regression models have been developed for a random sample consisting of: main, distribution and connection water pipelines and groups of pipes extracted from this random sample: steel pipes, gray cast iron pipes, DN80, DN100, DN150 , DN100 made of steel, DN100 made of gray cast iron, DN150 made of steel, pipes built in the following time intervals: 1885-1920, 1922-1938, 1939-1961, 1962-1978, 1979-1994, 1995-2012.

The analysis was performed in the Statistica software package (Statsoft) using multiple regression.

2.4. Characteristics of the analyzed water distribution subsystem

The water distribution subsystem is located in Upper Silesia (Poland) and is responsible for the supply of water to approximately 160 000 residents of the commune (163 thousand inhabitants in 2011 and 157 thousand in 2014). The source of water supply for the distribution subsystem are groundwater and surface water intakes owned by Górnośląskie Przedsiębiorstwo Wodociągów S. A. (GPW). The sup-

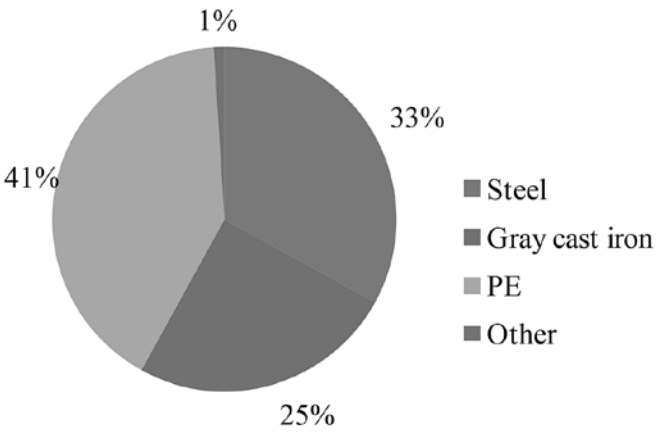


Fig. 2. The material composition of water supply system in 2014. On the graphs: "Other" – asbestos and PVC

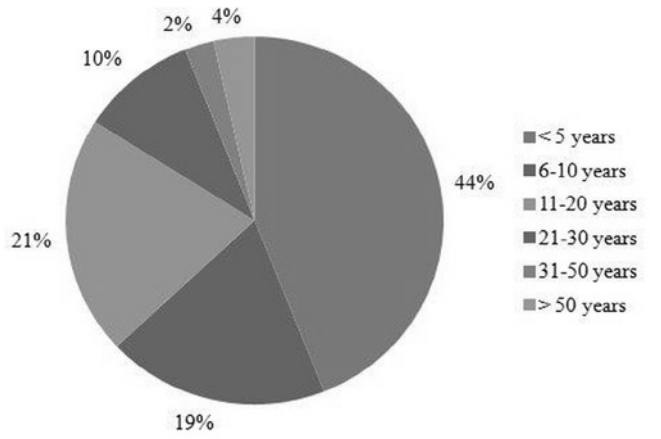


Fig. 3. The age structure of pipelines in 2014

ply of water from the GPW takes place through 56 chambers and on average in the analyzed years from 2011 to 2014 it was 7.3 million m^3/year, while water supply to residents amounted to average approximately 6.2 million m^3/year.

The water supply network in the diameter range DN20 - DN600 is approximately 450 km long, of which approximately 200 km is located in the mining area. Water supply pipelines are mainly made of: polyethylene, steel and gray cast iron (Fig. 2), and are mostly operated (60%) for less than 10 years (Fig. 3).

The value of the failure intensity indicator for the analyzed water distribution subsystem in particular years (Tab. 2) indicates a high failure rate of water pipelines $\lambda > 0.5$ damage/(km·year).

Table 2. Damage intensity for the water distribution subsystem in 2011-2014

Year	Number of failure	Lenght of the network; km	Unit failure intensity rate λ; damage/(km·year)
2011	445	443,9	1,00
2012	532	444,9	1,20
2013	299	456,4	0,66
2014	285	454,6	0,63

2.5. Characteristics of mining tremors occurring in the analyzed area

The analyzed water distribution subsystem is located in the commune, where active mining operations have been conducted since the 19th century.

At present, mining tremors are registered on 7 measurement stations equipped with Amax-GSI equipment, manufactured at the Central Mining Institute. The Amax-GSI apparatus [41] is a multi-channel specialized apparatus used for direct measurement of vibrations on the surface. The registration of earth vibration accelerations takes place using a recorder and a three-component piezoelectric accelerometer assembly. Thanks to the integration of the signal, a vibration velocity record is obtained in three perpendicular planes. The energy range recorded in the analyzed tremors region ranges from $1 \cdot 10^2$ to $1 \cdot 10^7$ J (Fig. 4).

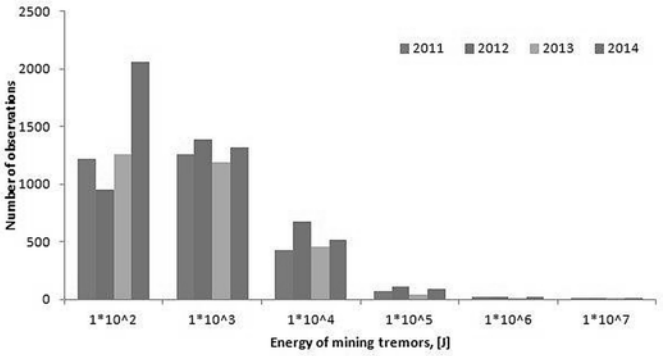


Fig. 4. Mining tremors registered in the mining area in the energy range from $1 \cdot 10^2$ to $1 \cdot 10^7$ J in 2011-2014.

In the analyzed area, the most tremors are weak, imperceptible by humans ($E < 10^4$ J) and registered only by specialist equipment.

As part of the study, only tremors with $E \geq 10^6$ J were recorded on 6 measuring stations due to the large number of incomplete recorded at one of the measuring station. Finally, 113 mining tremors (252 registrations) were adopted for the analysis, whose energy range is presented in Fig. 5.

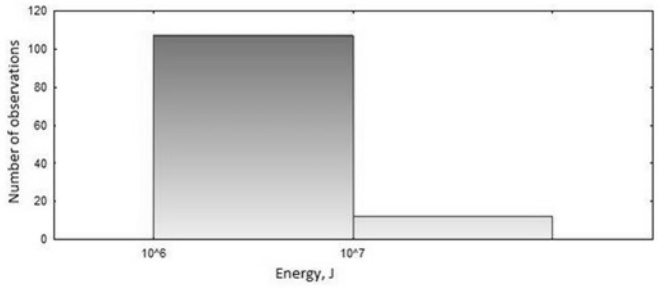


Fig. 5. Histogram of tremors energy occurred in the years 2010 - 2014.

Fig. 6 presents the values of the horizontal velocity component of the vibrations, recorded at measurement stations. The range of recorded values is from 0.0002 to 0.3150 m/s, while the median for all stations is similar and amounts to approximately. 0.0030 m/s. The highest value of ground vibration velocity was registered on 16.02.2013 at the 6th measurement station, while the smallest values at stations 5 and 2.

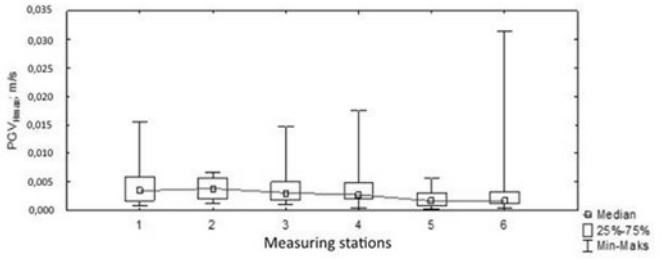


Fig. 6. Distribution of the medians of the horizontal component value of peak ground vibrations velocity for measurement stations for tremors with energy E ≥ 106 J occurred in 2010 - 2014.

3. Local attenuation relation

The estimation of the regression equation parameters was carried out using the least squares method. The model eliminated independent variables, which were characterized by the smallest value of the explained to unexplained variance. Finally, the local attenuation relation for logPGVHmax, for the significance level λ = 0.05, took the following form:

$$\log PGV_{Hmax} = 0{,}44937 \cdot \log E - 1{,}49304 \cdot \log R + 0{,}07623 \cdot S_i - 0{,}88661 \quad (8)$$

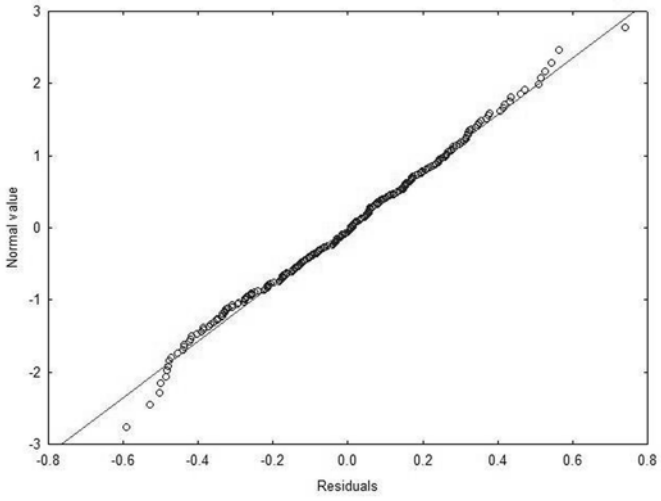


Fig. 7. Normal probability plot of residuals for a prediction model of $logPGV_{Hmax}$

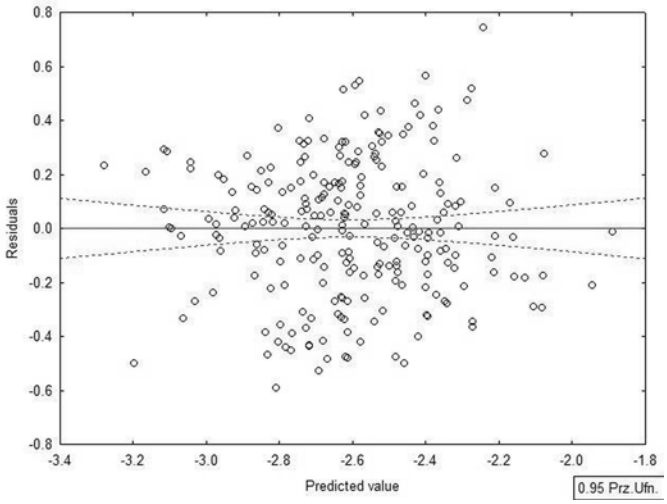


Fig. 8. Scatter plot of residuals for a prediction model of $logPGV_{Hmax}$

where:

PGVHmax – horizontal component of Peak Ground Velocity ; [m/s],
E – tremor energy, [J],
R – the distance from the source of the tremor for pseudo depth h = 1022m for which the value of the standard estimation error according to equation (3) assumes the smallest value, [m],
Si – quality variable taking value Si = 0 for a ground type B and Si = 1 for a ground type C, [-].

The correlation coefficient for the developed model is 0.68840. The value of the coefficient of determination $R^2 = 0.47390$ means a moderate fit of the model to the analyzed data. The value of coefficient of determination depends on: inaccuracy determination of energy and source of tremors [2], do not take into account the differentiation of the energy radiation from the source [2], not fully recognized geological structure [2], measurement errors of the accelerometers, regression model errors. In most of the seismic wave propagation models found in the literature [2,6,10,26,27], the coefficient of determination assumes similar values. Accordingly, it was decided to continue the analysis based on the obtained regression model.

The standard error of the estimate is Se = 0.25344 and means that the estimated values of the logPGVHmax variable differ from the empirical values by an average of 0.25344.

The construction of the multiple regression model was completed by residual analysis, which confirmed that the model meets the assumptions of the classic least squares method (Fig. 7, Fig. 8).

4. The influence of selected factors on the failure intensity of the water distribution subsystem

The results of the failure intensity analysis of all water supply network and groups of steel and gray cast iron pipes are presented in Tab. 3.

In the case of a groups of: all pipes and steel pipes, the obtained values of coefficients of determination indicate a good fitting of the models to the data. The PGV_{Hmax} parameter informing about the impact of mining tremors is an important variable for these models and has a positively affects on the dependent variables. The most significant influence on the dependent variable in the model for all pipes exerts successively: DN and M, Tp and PGV_{Hmax}, and in the case of steel pipes: DN and PGV_{Hmax}. For instance, for a model built for steel pipes if PGV_{Hmax} increase of 0.01 m/s, the dependent variable logλspvs will increase of 0.72906 (the damage intensity will increase of 5.35797 damage/(km·4 year)).

For a gray cast iron pipes, a weaker relationship was obtained ($R^2 = 0.47845$) and the variable informing about the impact of mining tremors on the water pipes turned out to be irrelevant.

In the whole range of analyzed pipe diameters (DN20 - DN600), only for DN80, DN100 and DN150 statistical relations between selected factors and the failure intensity of pipelines were obtained. For the remaining variables, either there was a lack of sufficient data or independent variables turned out to be irrelevant, or the correlation and determination coefficients were zero. A comparison of the obtained multiple regression models is contained in Tab. 3.

Only in the model built for a pipes with a diameter DN100, the variable describing the impact of mining tremors turned out to be significant. The model explains about 40% of the variability of the dependent variable. The correlation coefficient is close to 0.65. The following influence on the dependent variable has the following effects: Tp, PGV_{Hmax} and W. The models built for DN80 and DN150 have low values of coefficient of determination.

In the case of the analysis of failure intensity, taking into account diameters and materials from which the pipes were made, it was possible to build regression models only for pipes: DN100 made of steel,

Table 3. Regression models for the failure intensity of water pipelines.

Type of pipelines	Regression model	R, R^2, Se
All pipes	$\log\lambda spv = -0{,}03547 \cdot DN + 0{,}68996 \cdot M + 73{,}53762 \cdot PGV_{Hmax} + 0{,}00085 \cdot Tp + {,}62527$	R= 0,84113 R^2= 0,70750 Se = 1,11430
Steel pipes	$\log\lambda spvs = -0{,}04067 \cdot DN + 72{,}90571 \cdot PGV_{Hmax} + 3{,}17737$	R= 0,85774 R^2= 0,73572 Se = 1,03300
Gray cast iron pipes	$\log\lambda spvz = -0{,}00297 \cdot DN - 0{,}00068 \cdot Tp - 1{,}42276$	R= 0,69170 R^2= 0,47845 Se = 0,26376

Table 4. Regression models for the failure intensity of water pipelines with diameters DN80, DN100 and DN150.

Diameter	Regression models	R, R^2, Se
DN80	$\log\lambda sv80 = -0{,}01010 \cdot W - 1{,}20001$	R= 0,53986 R^2= 0,29144 Se = 0,39376
DN100	$\log\lambda sv100 = -0{,}03320 \cdot W + 67{,}93168 \cdot PGV_{Hmax} - 0{,}00070 \cdot Tp - 1{,}60510$	R= 0,64731 R^2= 0,41902 Se = 0,26561
DN150	$\log\lambda sv150 = -0{,}00113 \cdot Tp - 1{,}55594$	R= 0,66583 R^2= 0,44333 Se = 0,29573

Table 5. Regression models for the failure intensity for a water pipelines with diameters and materials

Diameter and material	Regression model	R, R^2, Se
DN100 steel	$\log\lambda sv100s = 64{,}44562 \cdot PGV_{Hmax} - 0{,}00083 \cdot Tp - 1{,}70361$	R= 0,68320 R^2= 0,46677 Se = 0,25054
DN100 grey cast iron	$\log\lambda sv100z = -0{,}00088 \cdot Tp - 1{,}63868$	R= 0,49176 R^2= 0,24183 Se = 0,27936
DN150 steel	$\log\lambda sv150s = -0{,}00135 \cdot Tp - 1{,}45395$	R= 0,74349 R^2= 0,55278 Se = 0,28652

Table 6. Regression models for the damage intensity of to water supply network pipes and connections built in different time periods.

Group	Year	Regression model	R, R^2, Se
1	1885-1920	$\log\lambda_{spv(1885\text{-}1920)} = -0{,}03320 \cdot DN + 1{,}34550 \cdot M + 216{,}44430 \cdot PGV_{Hmax} + 0{,}03000 \cdot P$	R= 0,90274 R^2= 0,81494 Se = 1,06430
2	1922-1938	$\log\lambda_{spv(1922\text{-}1938)} = -0{,}01687 \cdot DN + 1{,}61912$	R= 0,73161 R^2= 0,53525 Se = 1,57350
3	1939-1961	$\log\lambda_{spv(1939\text{-}1961)} = -0{,}04136 \cdot DN + 1{,}08442 \cdot M + 2{,}33101$	R= 0,86648 R^2= 0,75078 Se = 1,09150
4	1962-1978	$\log\lambda_{spv(1962\text{-}1978)} = -0{,}03291 \cdot DN + 0{,}62525 \cdot M + 50{,}00144 \cdot Tp + 1{,}28402$	R= 0,86310 R^2= 0,74495 Se = 0,99725
5	1979-1994	$\log\lambda_{spv(1979\text{-}1994)} = -0{,}04550 \cdot DN + 3{,}65622$	R= 0,87887 R^2= 0,77242 Se = 0,96855
6	1995-2012	$\log\lambda_{spv(1995\text{-}2012)} = -0{,}050443 \cdot DN - 0{,}645707 \cdot M + 4{,}442355$	R= 0,90993 R^2= 0,82797 Se = 0,73568

DN100 made of gray cast iron and DN150 made of steel. The results of the analyzes carried out are summarized in Tab. 4.

Only in the model for pipes made of steel with a diameter DN100, the variable describing the impact of mining tremors (PGV_{Hmax}) turned out to be an important variable. The variable has a positive effect on the failure intensity. In the model, also Tp and the intercept parameter have a significant impact on the dependent variable. The model explains about 47% of the variability of the $\log\lambda sv100s$.

For pipes made of gray cast iron with a diameter DN100 and for pipes made of steel with a diameter DN150, only Tp and the intercept parameter are an important variables. The dependence obtained for grey cast iron pipes DN100 is very weak, the value of the coefficient of determination is only 25% and indicates the lack of dependence. A better model was obtained for steel pipes DN150, in which the coefficient of determination is approximately. 55%.

Due to the fact, that in the research water pipelines at the age of even 129 were analized, it seems justified to perform analysis of failure rates of pipes made at particular time intervals (Tab. 6).

The obtained regression models are characterized by a very good fit to the observation, the coefficient of determination is in most cases in the range of 74 to 83%. Only the model for pipes built in 1922 - 1938 has a coefficient of determination around 53%. The variable characterizing mining tremors is significant only for the model did for pipes built in 1885 - 1920. According to the above, it can be concluded that the oldest analyzed pipelines may be least resistant to occurring mining tremors. In the presented model, an important variable was also the pressure in the place of pipe failure, which may also indicate that only the oldest pipes are less resistant to this factor. For pipes built in 1922 - 1938 and 1979 - 1994, only diameter with a intercept parameter had a significant impact on the failure intensity of the water pipes, which is most likely the result of significant differences in the way the pipes with each diameters were built.

For the remaining groups of pipes, analyzed due to the time of their construction, the parameters important in the created models were: diameter and

material of the pipes, operation time without failure and intercept parameter.

The analysis of the residues, for each regression models discussed above, showed that the assumptions of regression analysis in terms of residual normality, model linearity and homoscedasticity for the following variables are fulfilled: logλspvz, logλsv80, logλsv100, logλsv150, logλsv100s, logλsv100z, logλsv150s . In the case of other models, the conditions of the model's linearity with regard to parameters and homoscedasticity were not met. On the scatter plots obtained, two groups of points are visible, due to a large discrepancy between the value of failure of the water supply network and connections. The described situation indicates heterogeneity of the random sample, accepted for analysis and may affect the quality of the models obtained. However, this does not change the fact that the values of coefficient of determination obtained in the models had a good value, due to the greater amount of data taking into account the whole group of pipes. Accordingly, it seems reasonable to conduct a similar analysis individually for the water supply network and connection.

5. Summary and conclusions

The analyzes presented in the article are the first attempt of this type to determine the impact of mining tremors on the failure intensity of the water distribution subsystem.

The research was based only on historical data obtained from the real object, which made it impossible to repeat and eliminate some measurement errors. It should also be emphasized, that the research covered the commune with an area of 69.44 km^2. This results in the possibility of occurrence of many additional factors, not taken into account/unidentified in the conducted research, which could have influenced on the dependent variables, and thus also the quality of the models obtained.

The analyzes presented in the article allowed to formulate the following conclusions:

1. The independent variable (PGV_{Hmax}) informing about the impact of mining tremors on the failure intensity of the water distribution subsystem's pipelines turned out to be significant in the constructed regression models for selected groups of water distribution subsystems. This indicates the existence of a factor, which has not been accounted for so far, affecting the failure rate of water pipes, located in mining areas.
2. The independent variable, which in most of the built regression models had the biggest influence on the dependent variable, was the diameter of the pipes (9 out of 15 models) and sequentially the operation time without failure (7 out of 15 models) and PGV_{Hmax} and the material from which the pipes are made (5 out of 15 models).
3. The methodology of analysis and assessing the impact of mining tremors on the failure intensity of the water distribution subsystems can be applied to each subsystem of water distribution and other pressure - related linear underground facilities, such as: heating, gas, pressure sewerage, located on the impact mining tremors areas, provided that the following data will be available: failure rate of linear objects, the type of ground on which the objects are located, energy values of tremors, coordinates of the tremors epicenter and the local attenuation relations in the considered area or in the absence of such relations - PGV_{Hmax} values, as well as the coordinates of measuring stations and the type of ground on which they are located.
4. The developed regression models can be the basis for the development of a computer program that could be used to forecast the damage of the water distribution subsystem after the mining tremors, with defined parameters for instance, based on a mining tremors forecast, when starting operation in a new mining field.

References

1. Atkinson G. M., Boor D. M., Earthquake ground-motion prediction equation for Western North America, Bulletin of the Seismological Society of America 2006; 96: 2181-2205, https://doi.org/10.1785/0120050245.
2. Bańka P., Kołodziejczyk P., Lier E., Wykorzystanie wyników pomiarów parametrów drgań gruntu do wyznaczenia wartości współczynnika amplifikacji drgań, Przegląd Górniczy 2016, 4 (1121): 71-79.
3. Boron P., Dulińska J., The dynamic analysis of a steel pipeline under a seismic shock, Procedia Engineering 2017, 199: 104 – 109, https://doi.org/10.1016/j.proeng.2017.09.166.
4. Bouziou D., O'Rourke T. D., Response of the Christchurch water distribution system to the 22 February 2011 earthquake, Soil Dynamics and Earthquake Engineering 2017, 97: 14 – 24, https://doi.org/10.1016/j.soildyn.2017.01.035.
5. Bubtiena A. M., El Shafie A. H., Jaafar O., Performance improvement for pipe breakage prediction modeling using regression method, International Journal of the Physical Sciences 2011, 25 (6): 6025 – 6035, http://www.academicjournals.org/journal/IJPS/article-full-text-pdf/F7B3D0625658.
6. Chodacki J., New ground motion prediction equation for peak ground velocity and duration of ground motion for mining tremors in Upper Silesia, Acta Geophysica 2016, 64 (6): 2449 – 2470, http://agp.igf.edu.pl/files/64/6/Chodacki.pdf.
7. Clark R., Deininger R.A., Protecting the nation's critical infrastructure: the vulnerability of U.S. water supply systems, Journal of Contingencies and Crisis Management 2000, 8 (2): 73–80, doi: 10.1111/1468-5973.00126.
8. Council Directive 2008/114/EC of 8 December 2008 on the identification and designation of European critical infrastructures and the assessment of the reed to improve their protection, Official Journal of the European Union JOL_2008_345_R_0075_01.
9. Cubillo F., Pérez P., Water distribution system risk assessment method, Procedia Engineering 2014, 89: 355-362, https://doi.org/10.1016/j.proeng.2014.11.199.
10. Dubiński J., Mutke G., Tatara T., Muszyński L., Barański A., Kowal T., Zasady stosowania zweryfikowanej górniczej skali intensywności drgań GSIGZWKW-2012 do prognozy i oceny skutków oddziaływania wstrząsów indukowanych eksploatacją złóż węgla kamiennego w zakładach górniczych Kompanii Węglowej S. A. na obiekty budowlane i na ludzi, 2013 – instrukcja.
11. Dulińska J., Oddziaływanie drgań powierzchniowych wywołanych wstrząsami górniczymi w rejonie GZW I LGOM na konstrukcję gazociągu, Wstrząsy górnicze – charakterystyka parametrów drgań oraz kryteria oceny wpływu na obiekty budowlane, Główny Instytut Górnictwa, Katowice, 2010.
12. Dulinska J.M., Jasinska D., Performance of Steel Pipeline with Concrete Coating (Modeled with Concrete Damage Plasticity) under Seismic Wave Passage, Applied Mechanics and Materials 2014, 459: 608-613, doi:10.4028/www.scientific.net/AMM.459.608.
13. EN 1998 Eurocode 8: Design of structures for earthquake resistance.

14. Esposito S., Iervolino I.; PGA and PGV Spatial Correlation Models Based on European Multievent Datasets, Bulletin of the Seismological Society of America 2011, 101 (5): 2532-2541, https://doi.org/10.1785/0120110117.
15. Farmania R., Kakoudakisb K., Behzadianc K., Butler D., Pipe failure prediction in water distribution systems considering static and dynamic factors; Procedia Engineering 2017, 186: 117-126, https://doi.org/10.1016/j.proeng.2017.03.217.
16. Frej A., Zuberek W. M., Local effects in peak accelerations caused by mining tremors in bytom syncline region (Upper Silesia), Acta Geodynamica et Geomaterialia 2008, 5 (2): 115–122, https://www.irsm.cas.cz/materialy/acta_content/2008_02/3_Frej.pdf.
17. Gangl G., Fuchs-Hanusch D., Stadlober E., Kauch P.; Analysis of the failure behaviour of drinking water pipelines; Water Science and Technology: Water Supply 2007, 7 (5-6): 219-225, doi: 10.2166/wst.2011.507.
18. Golik, A. Mendecki, M., Ground-motion prediction equations for induced seismicity in the main anticline and main syncline, Upper Silesian Coal Basin, Poland 2012, 60 (2): 410 – 425, https://doi.org/10.2478/s11600-011-0070-9.
19. Hamdala .F. K., Sagar G. Y., Statistical analysis of pipe breaks in water distribution systems in Ethiopia, the case of Hawassa, IOSR Journal of Mathematics 2016; 12 (3): 127 – 136, http://www.iosrjournals.org/iosr-jm/papers/Vol12-issue3/Version-4/N120304127136.pdf.
20. Hotloś H., Ilościowa ocena wpływu wybranych czynników na parametry i koszty eksploatacji sieci wodociągowych, Wrocław, 2007.
21. Hotloś H., Mielcarzewicz E., Metody oceny udziału szkód górniczych w uszkadzalności sieci wodociągowych, Materiały konferencyjne Rola GPW w systemie zaopatrzenia w wodę dziś i jutro, Górnośląskie Przedsiębiorstwo Wodociągów w Katowicach i PZiTS o/Katowice, Ustroń, 1997.
22. Hotloś H., Mielcarzewicz E., Warunki i ocena niezawodności działania sieci wodociągowych i kanalizacyjnych na terenach górniczych, Monografia, Prace Naukowe Instytutu Inżynierii Ochrony Środowiska Politechniki Wrocławskiej, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 2011.
23. Isoyama R., Ishida E., Yune K., Shirozu R, Seismic damage estimation procedure for water supply pipeline, Proceedings of the twelfth world conference on earthquake engineering 2000, 1762, http://www.iitk.ac.in/nicee/wcee/article/1762.pdf.
24. Joachim K., Kalisz P., 2010: Awarie sieci gazowych na terenach górniczych. Główny Instytut Górnictwa „Górnictwo i Środowisko”2010, 4(1): 95- 105.
25. Kalisz P., Stec K., 2016: Oddziaływanie wstrząsów górniczych na gazociągi, Przegląd Górniczy 2016, 72(10): 1 – 8, http://www.sitg.pl/przegladgorniczy/spis-wydawniczy.html
26. Kurzeja J., Estymacja czasu trwania drgań gruntu generowanych silnymi wstrząsami w kopalniach GZW, Przegląd Górniczy 2016, 7 (1124): 51 – 56.
27. Kurzeja J., Seismometric monitoring in the area of the Piekary Śląskie junction of the A1 motorway in terms of recording the vibrations resulting from mining tremors, Journal of Sustainable Mining 2017, 16: 14 – 23, https://doi.org/10.1016/j.jsm.2017.06.002.
28. Kuś K. i inni, Podstawy projektowania układów i obiektów wodociągowych, Politechnika Śląska, Skrypt uczelniany nr 1854, Gliwice, 1995.
29. Kwietniewski M., Miszta-Kruk K., Piotrowska A., 2011: Wpływ temperatury wody w sieci wodociągowej na jej awaryjność w świetle eksploatacyjnych badań niezawodności, Wydawnictwo Politechniki Krakowskiej, Czasopismo Techniczne 1 – Ś/2011, 1 (108): 113-129.
30. Kwietniewski M., Rak J. Niezawodność infrastruktury wodociągowej i kanalizacyjnej w Polsce. Monografie Komitetu Inżynierii Lądowej i Wodnej PAN, Studia z Zakresu Inżynierii, 67, Warszawa 2010,
31. Kwietniewski M., Roman M., Kłoss-Trębaczkiewicz H.: Niezawodność wodociągów i kanalizacji. Arkady. Warszawa 1993
32. Lasocki, S. (2013), Site specific prediction equations for peak acceleration of ground motion due to earthquakes induced by underground mining in Legnica-Głogów Copper District in Poland, Acta Geophysica 2013, 61(5): 1130-1155, https://doi.org/10.2478/s11600-013-0139-8.
33. Lee D. H., Kim B. H., Lee H., Kong, J. S., 2009: Seismic behavior of a buried gas pipeline under earthquake excitations, Engineering Structures 2009, 31: 1011–1023, https://doi.org/10.1016/j.engstruct.2008.12.012
34. Mahmoodian M., Aryai V., Structural failure assessment of buried steel water pipes subject to corrosive environment, Urban Water Journal 2017 14 (10): 1023 – 1030, http://dx.doi.org/10.1080/1573062X.2017.1325500.
35. Mora-Rodríguez J., Delgado - Galván X., Ramos H. M., López-Jiménez P. A., An overview of leaks and intrusion for different pipe materials and failures, Urban Water Journal 2014, 11 (1): 1-10, http://dx.doi.org/10.1080/1573062X.2012.739630.
36. O’Rourke, M. J., Ayala, G., Pipeline damage due to wave propagation, Journal Geotechnical Engineering 1993, 119: 1490–1498, https://doi.org/10.1061/(ASCE)0733-9410(1993)119:9(1490)
37. O’Rourke M. J., Liu X., Response of buried pipelines subject to earthquake effects, Monograph no. 3, Multidisciplinary Center for Earthquake Engineering Research 1999, 33–57.
38. O’Rourke T. D., Jung J. K., Argyrou C., Underground pipeline response to earthquake-induced ground deformation, Soil Dynamics and Earthquake Engineering 2016, 91: 272 – 283, https://doi.org/10.1016/j.soildyn.2016.09.008
39. O’Rourke T. D., Toprak S., Sano Y., Factors Affecting Water Supply Damage Caused by the Northridge Earthquake, Proceedings of 6th US National Conference on Earthquake Engineering 1998: 1–12.
40. Pietrucha-Urbanik K, Studziński A., Case study of failure simulation of pipelines conducted in chosen water supply system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 317–323, http://dx.doi.org/10.17531/ein.2017.3.1.
41. Pilch R., Szybka J., Tuszyńska A., Application of factoring and time-space simulation methods for assessment of the reliability of water-pipe networks, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (2): 253-258, http://www.ein.org.pl/sites/default/files/2014-02-12.pdf.
42. Pineda-Porras O., Ordaz M., A new seismic intensity parameter to estimate damage in buried pipeline due to seismic wave propagation, Journal of Earthquake Engineering 2007, 11: 773–786, http://dx.doi.org/10.1080/13632460701242781.
43. Rak J., Bezpieczeństwo systemów zaopatrzenia w wodę, Polska Akademia Nauk, Instytut Badań Systemowych, Warszawa 2009.
44. Rak J., Wybrane aspekty bezpieczeństwa systemów wodociągowych, Oficyna Wydawnicza Politechniki Rzeszowskie 2015.
45. Rak J., Tchurzewska-Cieślak B., Studziński A., Pietrucha-Urbanik K., Boryczko K., Niezawodność i bezpieczeństwo systemów zbiorowego zaopatrzenia w wodę, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2012.
46. Rezaei H., Ryan B., Stoianov I.; Pipe failure analysis and impact of dynamic hydraulic conditions in water supply networks; Procedia Engineering 2015, 119: 253 – 262, https://doi.org/10.1016/j.proeng.2015.08.883.

47. Scheidegger A, Leitão J P, Scholten L., Statistical failure models for water distribution pipes – A review from a unified perspective, Water Research 2015, 83: 237–247, https://doi.org/10.1016/j.watres.2015.06.027.
48. Si H., Midorikava S., New attenuation relations for peak ground acceleration and velocity considering effects of faulty type and site condition, Journal of structural and construction engineering 1999, 64 (523): 63 – 70, http://doi.org/10.3130/aijs.64.63_2.
49. Takada S. and Tanabe K., 1987: Three-dimensional seismic response analysis of buried continuous or jointed pipelines, Journal of Pressure Vessel Technology 1987, 109: 80–87, doi:10.1115/1.3264859.
50. Ustawa z dnia 26 kwietnia 2007 r. o zarządzaniu kryzysowym. Dz. U. 2007 Nr 89 poz. 590 wraz z póź. zm.
51. Ustawa z dnia 7 czerwca 2001 r. o zbiorowym zaopatrzeniu w wodę i zbiorowym odprowadzaniu ścieków Dz. U. 2017 poz. 328 wraz z póź. zm.
52. Wang, L. R.-L., Cheng, K.-M., Seismic response behavior of buried pipelines, Journal of Pressure Vessel Technology 1979, 101: 21–30, doi:10.1115/1.3454594.
53. Wieczysty A.: Niezawodność systemów wodociągowych i kanalizacyjnych Cz. I i II, Teoria niezawodności i jej zastosowania, Wydawnictwo Politechniki Krakowskiej, Kraków 1990.
54. Zasada działania rejestratora drgań AMAX-GSI, Instrukcja obsługi, Dokumentacja techniczno - ruchowa, Labolatorium Sejsmologii i sejsmiki górniczej.
55. Zembaty Z., Rockburst induced ground motion – a comparative study, Soil Dynamics and Earthquake Engineering 2004, 24 (1): 11 – 23, https://doi.org/10.1016/j.soildyn.2003.10.001.

Ewelina KILIAN-BŁAŻEJEWSKA
Institute of Water and Wastewater Engineering
Faculty of Energy and Environmental Engineering
Silesian University of Technology
ul. Akademicka 2A, 44-100 Gliwice, Poland

E-mail: eewelina.kilian@gmail.com

Article citation info:
WEREMCZUK A, BOROWIEC M, RUDZIK M, RUSINEK R. Stable and unstable milling process for nickel superalloy as observed by recurrence plots and multiscale entropy. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 318–326, http://dx.doi.org/10.17531/ein.2018.2.19.

Andrzej WEREMCZUK
Marek BOROWIEC
Michał RUDZIK
Rafał RUSINEK

STABLE AND UNSTABLE MILLING PROCESS FOR NICKEL SUPERALLOY AS OBSERVED BY RECURRENCE PLOTS AND MULTISCALE ENTROPY

STABILNY I NIESTABILNY PROCES FREZOWANIA SUPER STOPU NIKLU OBSERWOWANY Z WYKORZYSTANIEM WYKRESÓW REKURENCYJNYCH I ENTROPII WIELOSKALOWEJ

This paper discusses the stability of high-speed machining processes. The problem of harmful vibrations can usually be detected based on measured signal forces. Nevertheless, the chatter effect may be unrevealed and hence some alternative approaches of signal monitoring must be taken to detect it. In the discussed case of machining, process stability is determined by means of stability diagrams. The measured milling force components are investigated by various signal analysis methods. In addition to this, the analysis also uses recurrence plots, recurrence quantifications, composite multi-scale-entropy and as well the statistical approach. Results obtained by the different methods are presented and discussed.

***Keywords**: machining, recurrence plot, entropy, chatter.*

W niniejszym artykule omówiono stabilność procesu obróbki szybkościowej. Problem szkodliwych drgań zwykle może zostać wykryty na podstawie sygnału mierzonych sił. Niemniej jednak drgania samowzbudne mogą nie ujawnić się w sposób wyraźny, a niekiedy do ich wykrycia potrzebne jest alternatywne podejście do monitorowania sygnału. W przedstawionym procesie obróbki stabilność procesu oszacowano za pomocą wykresów stabilności. Zmierzone siły frezowania badano różnymi metodami analizy sygnału. W analizie wykorzystano wykresy rekurencyjne, wskaźniki rekurencyjne, entropię wieloskalową, a także podejście statystyczne. Przedstawiono wyniki różnych metod i omówiono ich porównanie.

***Słowa kluczowe**: obróbka skrawaniem, wykresy rekurencyjne, entropia, drgania samowzbudne.*

1. Introduction

Nickel alloys are often used in different branches of industry, from medical to aerospace industries. Nickel-based alloys are known to have very good strength and temperature resistance, therefore they are called superalloys. Due to their strength properties, superalloys are very difficult to machine. Additionally, the cutting of these materials can generate harmful vibrations during machining. This poses a serious problem for engineers. Undesired relative vibrations between the tool and the workpiece may deteriorate the quality of machined surfaces or even damage the machine tool and the workpiece. As a result, the cutting forces that depend on the tool geometry, material properties, feed rate and cutting speed can have a large amplitude, which leads to faster tool wear. These vibrations of the tool are known as chatter [39, 40]. In order to use the full capacity of a new fast machine and to achieve a potentially high material removal rate together with the desired surface quality, optimum machining parameters are necessary. The fundamental parameters which may improve efficiency of the cutting process include cutting depth and velocity. Usually, the selection of cutting depth and spindle rotational speed is made via Stability Lobes Diagram (SLD) which can be applied in high-speed machining (HSM) processes to optimize the maximum depth of cut at the highest available spindle speed. When the cutting depth exceeds the critical value, chatter vibrations can occur at some spindle speed, whereas if the cutting depth is below the critical value, the process is stable regardless of the spindle speed. Generally, in practice, the selection of optimal speed and depth of cut is difficult. Classical SLDs can be obtained by modal analysis of the tool-spindle system; nevertheless, in many research papers the milling process is first modelled and then the numerical results are used to determine SLD (for example [11, 15]. An alternative solution is to calculate stability lobes directly from delay-differential equations [22]. However, only a few papers report complete experimental verification of these stability lobes [24]. The literature reports numerous analytical, numerical and experimental methods for cutting stability prediction. For instance, Altintas and Budak [2] describe an analytical method for predicting milling stability lobes based on a mean of the Fourier series of dynamic milling coefficients. This method is fast but cannot predict additional stability regions and period doubling bifurcation for a low radial depth of cut [8]. To overcome this problem, a multi-frequency solution of chatter stability was developed by Budak and Altintas [3] and then extended by Merdol and Altintas [3, 27]. One of the most popular numerical methods for chatter prediction is the Finite Element Method (FEM) [1, 10, 23, 38]. Bayly et al. [4] propose the use of a temporal finite element analysis for milling and interrupted turning [13]. Moreover, Voronov [38] and especially Adetoro et al. [1] propose an improved model of the classical milling process. This new

model takes into account the well-known model and combines it with considerations about the nonlinearity of cutting force coefficients and the axial immersion angle as a function of the axial depth of cut. The model is validated and the theoretical findings show a very good agreement with the experimental results [1]. Although experimental findings are very often compared with FEM results [13], the problem of experimental signal analysis remains an open question. Various methods of signal analysis are applied to recognize chatter vibrations in cutting operations, including multifractal and wavelet approaches [18], multiscale entropy [20], Hilbert Huang transform [32, 19], recurrence analysis [33, 31], flicker-noise spectroscopy [21] and audio signal analysis [38]. In [7] the delay space reconstruction method is applied to show that the workpiece motion is characterized by fractal geometry. The auto-bispectra suggest a quadratic phase coupling among the spectral peaks associated with the cutter frequency. Finally, the authors propose a mechanics-based model with impact to explain the obtained results. Their predictions agree well with the experimental observations. Other researchers have noticed that the dynamic behaviour of the tool-workpiece system depends on the tool position in the workpiece. This phenomenon and the problem of the influence of several modes on stability lobes is discussed in [37]. On the other hand, modal interaction can affect surface roughness, particularly in the milling of thin-walled structures [34].

Numerical analysis of the cutting process is also useful for stability analysis. However, the models of cutting processes are described by delay differential equations, which may pose some problems in numerical calculations. To accelerate the integration procedure and overcome the difficulties, semi- or full-discretization methods for predicting milling stability are proposed in several papers. The semi-discretization method developed by Insperger and Stepan is an efficient numerical technique for stability analysis of linear delayed systems. Nonetheless, Ding et al. [8] propose a more effective full-discretization method for prediction of milling stability. Insperger [12] proposes the so-called act-and-wait concept for continuous-time control systems with feedback delay.

In the literature one can find a number of publications with practical approach to the problem of cutting efficiency. Wojciechowski et. al. [42] propose a method ensuring the reduction of forces and the improvement of efficiency in the finish ball end milling of hardened 55NiCrMoV6 steel. The primary objective of the paper is optimal selection of milling parameters (cutting speed and surface inclination angle), as this will ensure minimal cutting force values and increased process efficiency at the same time. From a practical point of view, surface roughness is very often analysed. Roughness depends on different factors including feed direction, axial and radial run-out errors, and cutting tool geometry. In the paper [45] an algorithm considering the effects of static and dynamic factors on surface roughness is proposed. A new approach to surface roughness parameters estimation during finish cylindrical end milling is presented in [41]. The proposed model takes account of the influence of cutting parameters, the tool's static runout and dynamic phenomena related to instantaneous tool deflection.

An interesting contribution to the problem is a numerical and experimental study of the dynamics of flank milling operations at low cutting rates presented in [28].The paper focuses on both the properties of cutting vibratory phenomena and their impact on the roughness of the machined surface. The study is based on a one-degree-of-freedom model of a mechanical machining system. The system consists of a rigid cutter and a flexible workpiece. The cutting force model is based on a regenerative mechanism. The roughness of the surface machined at high speed revolutions is studied for both forced vibrations occurring during stable cutting and self-excited vibrations occurring during unstable cutting. It is shown that forced vibrations have only a very slight impact on product roughness, while self-excited vibrations lead to a significant increase in roughness.

This paper contains the extended research on Inconel milling stability for various cutting speeds that correspond both to a stable and an unstable regions in SLD. A bit similar experiment performed for increasing the depth of cut at constant cutting speed is published in [16]. The main aim of the present paper is to investigate the effect of cutting speeds on cutting stability via multiscale entropy and the recurrence quantification technique. Moreover, the analysis of cutting forces is made with new recurrence plot quantifiers together with statistical indicators. The determination of proper stability indicators, irrespective of the rotational speed of the spindle (cutting velocity), is the main objective of the paper. The indicators could be applied in the future to build a chatter control system for detecting stability loss symptoms on the basis of statistical parameters, entropy or recurrence plot analysis.

The paper is organized as follows: section 2 presents the methodology of experimental research. Next, in section 3, a statistical analysis of cutting forces is performed, and then a recurrence analysis is presented in section 4. In section 5, the multiscale entropy method is applied to analyse milling process stability. Finally, section 6 contains conclusions.

2. Experiment and methodology

The experimental investigations were conducted on Inconel 718 cut on the Blue Bird MG6037PKK milling machine. The experimental setup, shown schematically in Fig.1, is composed of two subsystems: a modal analysis system and a dynamometer system.

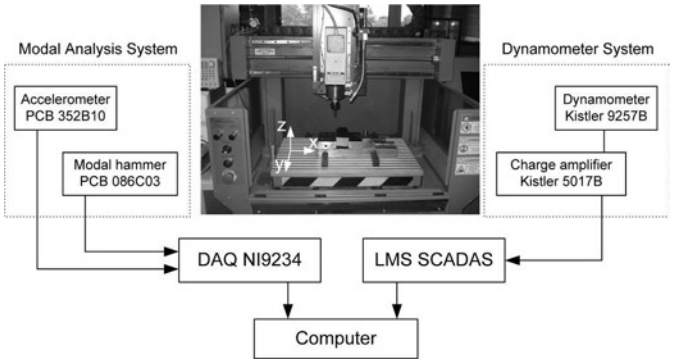


Fig. 1. Experimental setup of CNC milling machine with acquisition system scheme

The former is used to measure tool-holder stiffness and damping coefficient (modal parameters). It consists of the PCB 086C03 modal hammer, PCB 352B10 accelerometer and NI9234 data acquisition card (DAQ). The latter is used to measure the cutting force components (F_x, F_y and F_z) with the Kistler 9257B piezoelectric dynamometer connected to the Kistler 5017B signal conditioner and the SCADAS Mobile LMS analyser. Both experimental rigs are integrated with the computer system. Measurements are conducted in two steps. First, an impact test is performed to obtain data for a stability lobes diagram (SLD). The modal hammer is used to excite the tool, and then the resulting vibrations are measured by the low mass accelerometer mounted on the tool tip. Next, modal parameters in the form of frequency response function (FRF) are implemented to the CutPro9 software to calculate and plot an SLD (Fig. 2a). In the second step of the experiment, the unstable lobes are verified for a series of the spindle speed and the depth of cut marked as the points in Fig. 2b. The test is performed on Inconel 718 by a 12 mm diameter end milling cutter with flutes, made of PCD (FENES DIN 6527-A 12 KNZ4 13). The radial depth of cut equals 12mm (slot milling), the feed per flute is 0.01mm. The applied milling parameters are listed in Tables 1 and 2.

Table 1. Milling parameters applied in the experiment

Parameter	Value
Radial depth of cut (a_e)	12mm
Feed per flute (f)	0.01 mm
Axial depth of cut (a_p)	0.05 – 0.2 mm
Rotational speed of spindle (n)	3000 – 4150 rpm

Table 2. Milling parameters of measured points

Point name	n [rpm]	a_p [mm]	Point name	n [rpm]	a_p [mm]	Point name	n [rpm]	a_p [mm]
n1a1	3000	S0.05	n2a1	3500	0.05	n3a1	4150	0.05
n1a2	3000	0.10	n2a2	3500	0.10	n3a2	4150	0.10
n1a3	3000	0.15	n2a3	3500	0.15	n3a3	4150	0.15
n1a4	3000	0.20	n2a4	3500	0.20	n3a4	4150	0.20

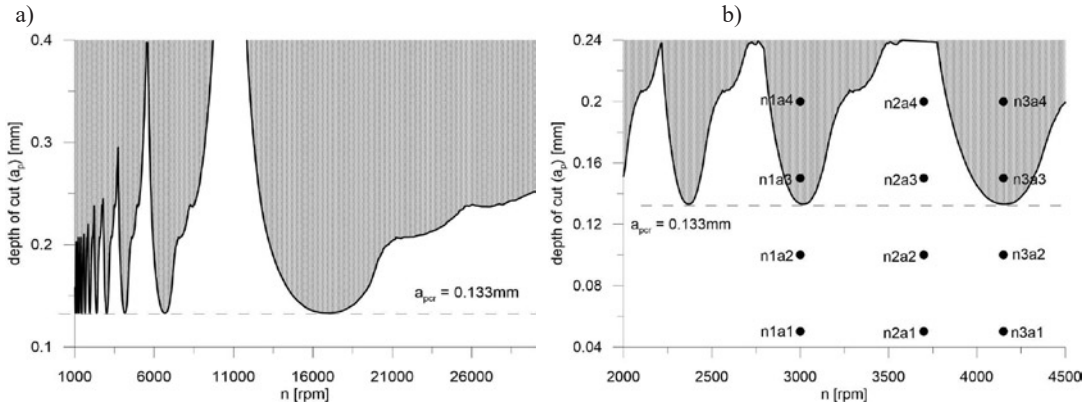


Fig. 2. Stability lobes diagram for Inconel 718, a_e=12mm, f=0.01mm.

The tool was changed after each test to provide identical cutting conditions and prevent tool wear which could affect process dynamics.

During the milling process, the forces F_x, F_y and F_z are recorded with a sampling rate of 2 kHz; this value is a necessary minimum because the natural frequency of the spindle-tool system is about 740 Hz. On the one hand, this sampling rate meets the Nyquist-Shannon sampling theorem, and on the other hand, it is low enough to record the milling process for a sufficient period of time. In addition to this, the presence of very long time series poses difficulty in a recurrence analysis. In order to avoid the aliasing phenomenon, the Kistler measuring system is provided with an anti-aliasing filter. Stable cutting occurs in the region below the stability boundary, while unstable machining should occur above the lobes (Fig. 2). According to the diagram, the critical cutting depth a_{pcr} is 0.133mm. For $a_p<a_{pcr}$ the process should be stable all the time regardless of the spindle speed. The subsequent section contains a force signal analysis aimed at analysing whether the cutting process with the assumed depth of cut and spindle speed is stable or not. This will help determine stability indicators taken directly from the experiment.

3. Statistical analysis

The statistical analysis is performed for three directions of the cutting force: F_x, F_y and F_z. Although F_x is the most important because its direction is compatible with the feed direction, all the three components are marked in Fig. 3 illustrating the distribution of the maximum cutting force. A typical behaviour can be observed, namely, with increasing the depth of cut the force components increase too. In the case of unstable cutting (n1a3, n1a4, n2a3, n2a4, n3a3, n3a4 in Fig.2b), the maximum forces rapidly increase to high values (see a_p=*0.15* and a_p=*0.2* in Fig.3). Examining the distribution of the standard deviation (Fig.4) one can observe that the greatest dispersion of the results occurs in the case of unstable points *n=3000rpm*, a_p=*0.2mm* (n1a4) and *n=4150rpm*, a_p=*0.2mm* (n3a4). For the points located in the stable area, the dispersion of the results is reduced. A large dispersion of the results for the force component F_y from the stable area for the spindle speed *n=3700rpm* can also be observed. Moreover, an analysis of the mean value is performed via kurtosis (Fig. 5). In the case of point n1a1, the kurtosis of all force components (F_x, F_y, F_z) proves that the force distribution is close to the normal one. The analysis of the distribution of the individual force components around the mean value demonstrates that the largest concentration of the results was obtained for point n1a3 (*n*=3000rpm, a_p=0.15mm).

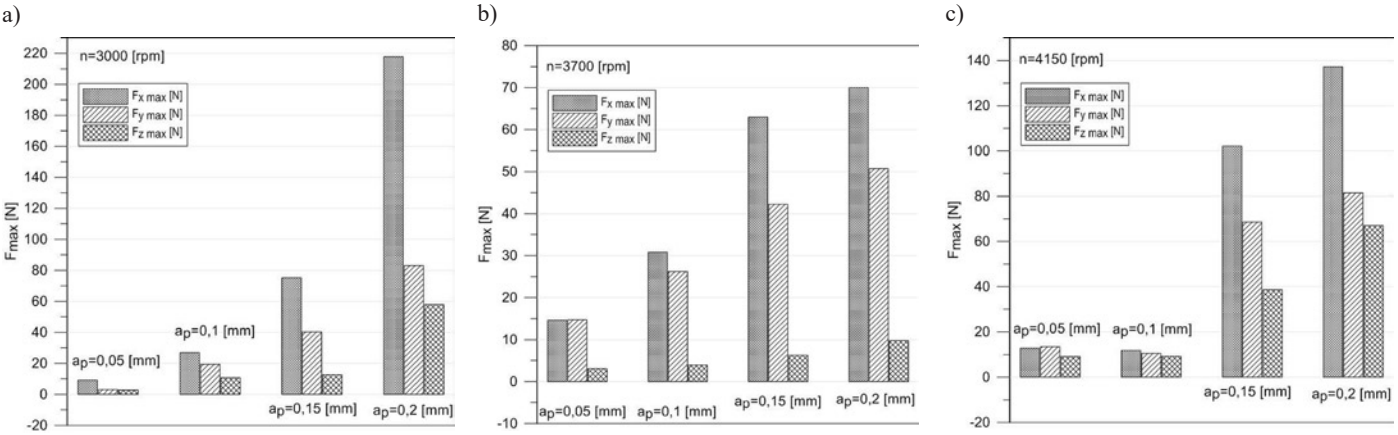


Fig. 3. Maximum values of cutting forces

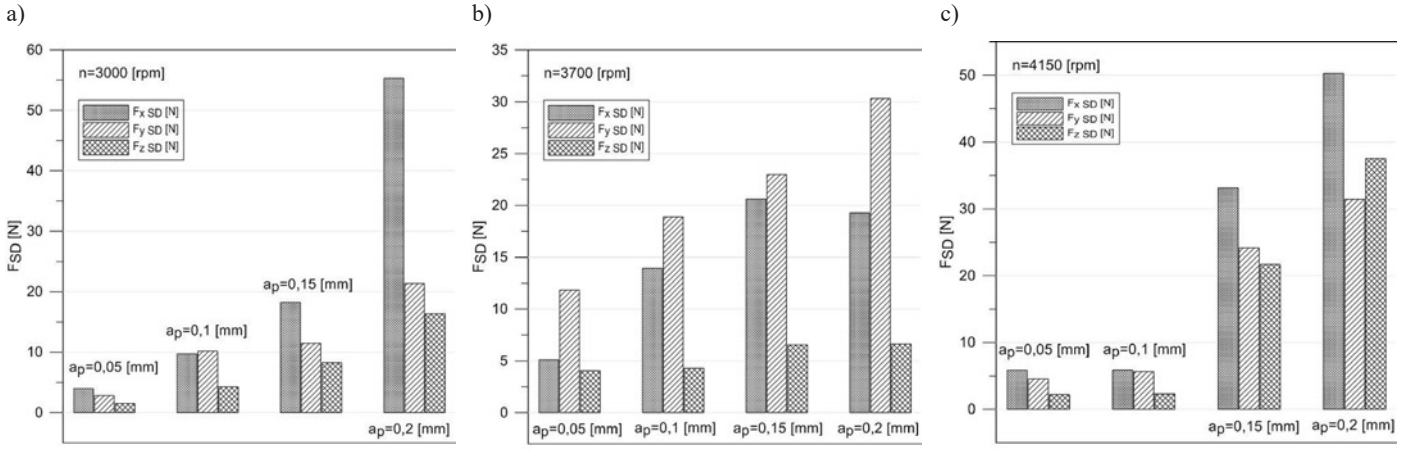


Fig. 4. Standard deviation of cutting forces

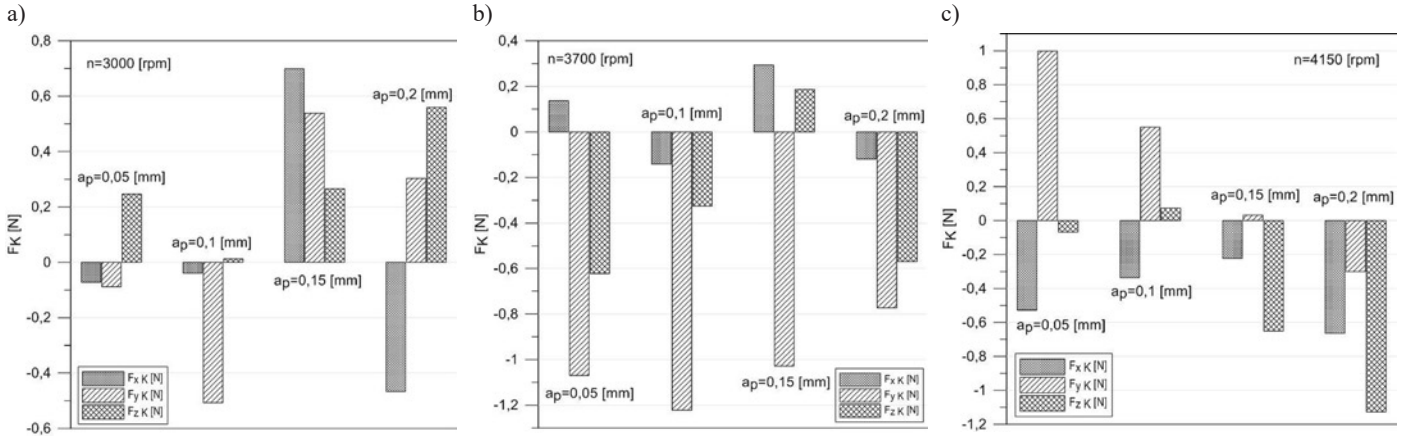


Fig. 5. Kurtosis of cutting forces

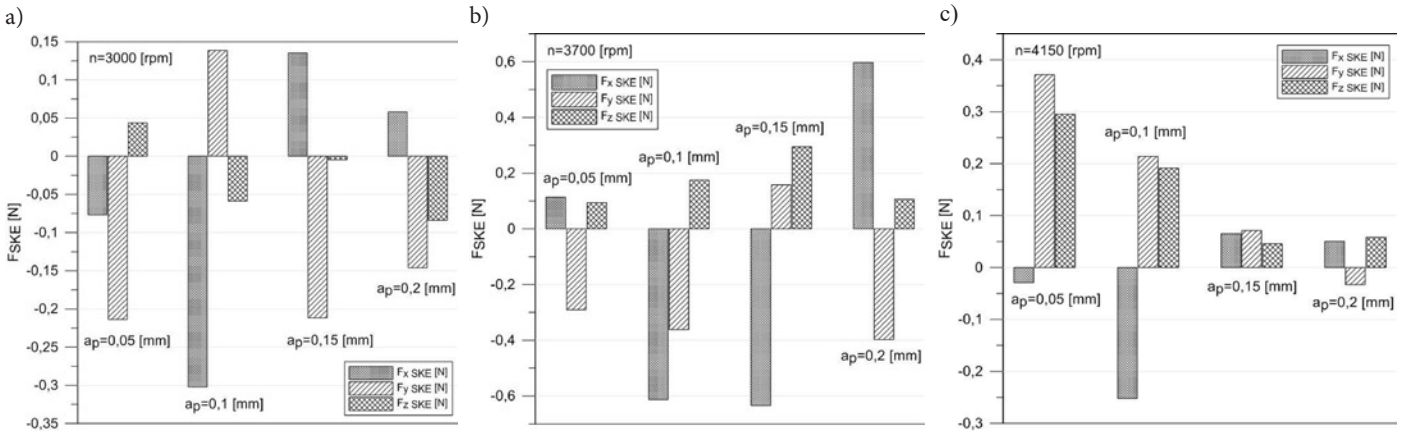


Fig. 6. Skewness of cutting forces

The results of all tests (particularly, the *y* component) at the spindle speed of *n=3700rpm* show a great deviation from the average value. Considering all analysed force signals, the smallest asymmetry of results is obtained for the unstable points n3a3 and n3a4 which correspond to *n=4150rpm*, a_p*=0.15mm* and a_p*=0.2mm*, respectively (Fig.6).

4. Recurrence plot

In order to ensure a more detailed force signals analysis and thorough verification of stability regions, the recurrence plot *(RP)* technique is employed. The *RP* approach provides a qualitative interpretation of hidden patterns of dynamical systems based on phase space reconstruction. This technique was introduced by Eckmann et al. [9]. The general idea of phase space reconstruction rests on the assumption that any time series x_i can be presented as a delayed vector S_i in an *m*-dimensional space called reconstructed phase space:

$$S_i = \left(x_i, x_{i+d}, x_{i+2d}, ..., x_{i+(\mathrm{m}-1)d}\right) \tag{1}$$

where *m* is the embedding dimension and *d* stands for the time delay.

Usually the embedding parameters, i.e., the time delay *(d)* and the embedding dimension *(m)*, can be estimated via the average mutual information function *(AMI)* and the false nearest neighbours method

(FNN). More information about the *AMI* and *FNN* functions can be found in [14, 29, 30]. In this study, the Tisean software [44] is used to obtain the embedding parameters. The recurrence plot technique is a graphical method designed to locate recurring patterns, non-stationarity and structural changes. It shows all time instants at which a state of the dynamical system recurs. The recurrence plot can be expressed as the matrix:

$$M_{ij} = \theta\left(\varepsilon - \left|s_i - s_j\right|\right) \tag{2}$$

where Θ is the Heaviside step function, ε is a tolerance parameter (threshold), s_i and s_j are delay vectors.

If the trajectory in the reconstructed phase space returns at the time i into the neighbourhood of ε where it was before at the time j then $M_{ij}=1$, otherwise $M_{ij}=0$. These results are plotted as black $(M_{ij}=1)$ and white $(M_{ij}=0)$ dots, respectively. From a practical point of view, *RPs* can be presented in a more useful and certainly more convenient form via recurrence quantification analysis *(RQA)*. *RQA* is a method which quantifies the number and duration of recurrences of a dynamical system presented by its state space trajectory. The measures of *RQA* were elaborated by Zbilut and Webber [44], and then developed and introduced to MATLAB by Marwan [25, 26]. The measures of *RQA* are based on the recurrence point density and the diagonal and vertical line structures of the recurrence plot. In this paper only the determinism *(DET)*, average length of the diagonal line *(L)* and *L* entropy *(Lent)* are applied to analyse milling stability. These measures are defined as follows: determinism *(DET)* is the fraction of recurrence points forming diagonal lines:

$$DET = \frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{i,j=1}^{N} M_{i,j}} \tag{3}$$

The average length of the diagonal lines *(L)*:

$$L = \frac{\sum_{l=l_{\min}}^{N} lP(l)}{\sum_{l=l_{\min}}^{N} P(l)} \tag{4}$$

L-entropy (Lent) is the Shannons entropy of the diagonal line segment distribution:

$$L_{ent} = -\sum_{l=l_{\min}}^{N} P(l)\ln P(l) \tag{5}$$

P(l) is the histogram of the lengths *l* of the diagonal lines, *P(v)* is the histogram of the lengths *v* of the vertical lines, *N* denotes the number of points on the phase space trajectory.

The recurrence quantification analysis can provide useful information even for short and non-stationary data, where other methods fail. *RQA* can be applied for various kinds of data to recognize dynamical behaviour. The F_x component is chosen for the analysis because it is naturally dependent on cutting parameters. The idea of phase space reconstruction (necessary for RP) assumes that any time series from

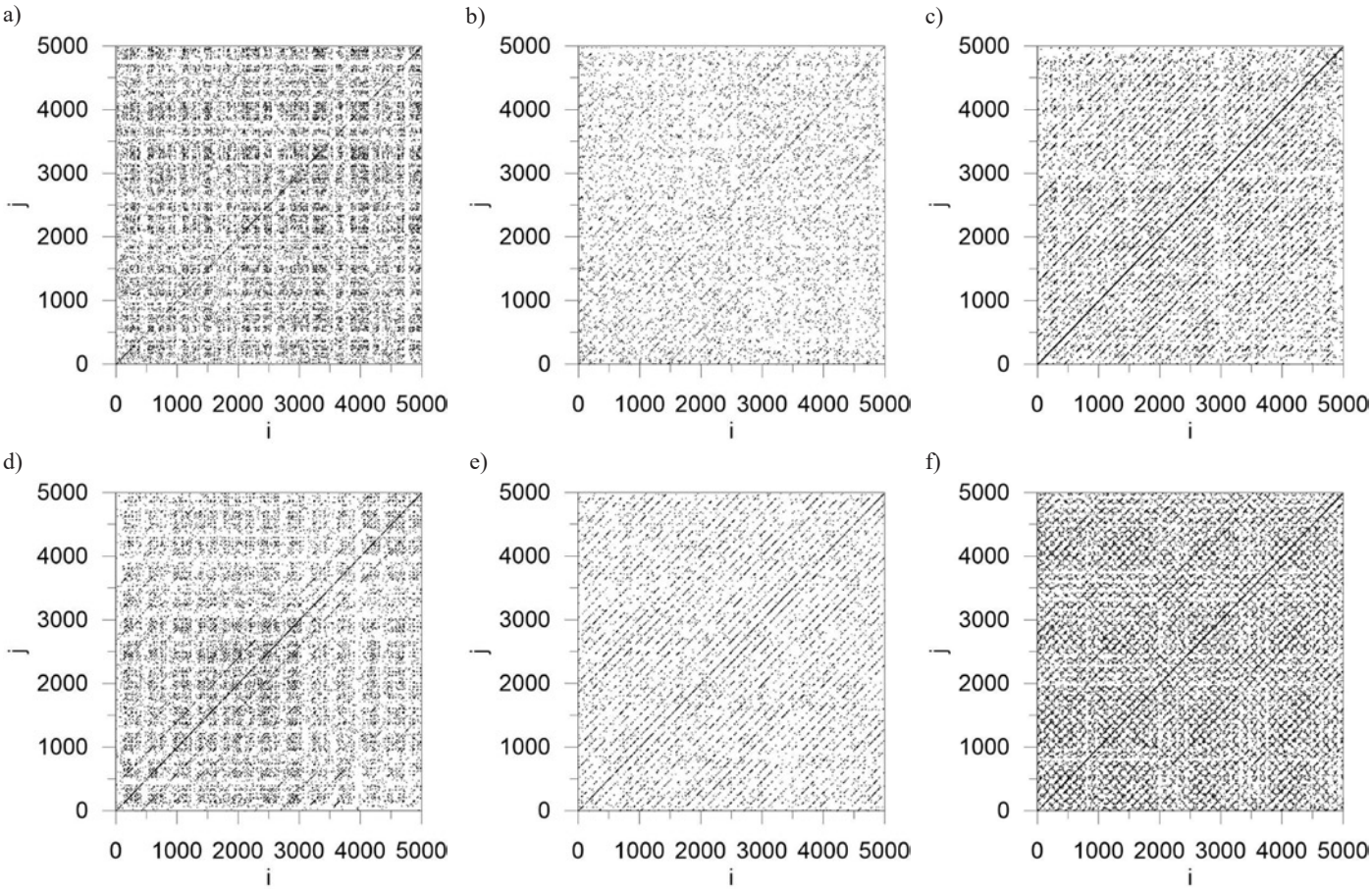


Fig. 7. Recurrence plots for cases of stable cutting (a)-n1a1, (b)-n2a1, (c)-n3a1, and unstable cutting (d)-n1a4, (e)-n2a4, (f)-n3a4.

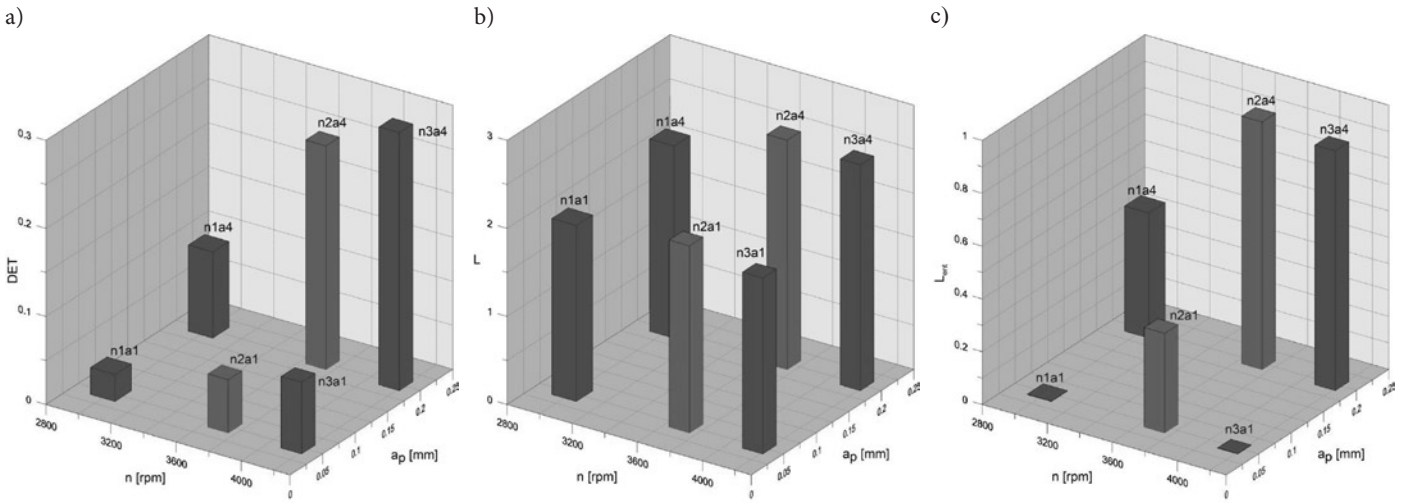


Fig. 8. Recurrence quantification analysis, determinism-DET (a), average length of the diagonal line-L (b) and L entropy-Lent (c).

an analysed process has the same part of information. Moreover, the amplitude of F_x is the biggest therefore a ratio of noise to signal is the smallest. The recurrence plots of the force (F_x) for the selected stable cutting cases (n1a1, n2a1, n2a4, n3a1) and unstable cutting cases (n1a4, n3a4) are presented in Fig. 7. Although relatively small differences can be noticed between the stable and unstable cases, it is rather difficult to spot differences between the recurrence patterns obtained for stable cutting (Fig. 7a,b,c,e) and unstable cutting (Fig. 7d,f); therefore, the recurrence quantification analysis is applied here (Fig. 8). *RQA* gives more distinct information in the form of index about recurrence and system dynamics. The determinism *(DET)* presented in Fig. 8a illustrates well the differences between the depth of cut *(a_p)* and the rotational speed *(n)*; however, the stable (n1a1, n3a1) and unstable (n1a4, n3a4) points can only be distinguished when *n=3000rpm* and *n=4150rpm*. For *n=3700rpm*, both points (n2a1, n2a4) are stable, therefore they should have a similar *DET* factor. Unfortunately, the *DET* is not able to find any differences in stability. Nonetheless, the *DET* factor points to an increase in the rotational speed *(n)*.

The average length of the diagonal lines *(L)* presented in Fig. 8b is the least significant stability factor. Moreover, the impact of the cutting depth *(a_p)* and the rotational speed *(n)* on *L* is insufficient. However, the Shannon's entropy (Lent) is also an efficient method for identifying cutting instability (Fig. 8c). Small values of Lent in the cases of stable cutting (n1a1, n3a1) considerably increase in the case of unstable cutting (n1a4, n3a4) when chatter occurs. Even in the case of *n=3700rpm* (green bar), when both points n2a1 and n2a4 are stable, the Lent shows an increase that is smaller than for *n=3000rpm* and *n=4100rpm*. This is because the increase is only caused by the change of the depth of cut and not by the change of stability.

Summing up, the *DET* shows the change of the cutting depth *(a_p)* and the rotational speed *(n)*, while the Lent is additionally able to distinguish stable and unstable cutting for various rotational speeds. A new cutting stability indicator can be proposed which is defined as a ratio of the Lent for unstable to stable points *(LentR)*. This ratio is vital because for *n=3000rpm LentR=471*, for *n=4150rpm LentR=908*, while in the case of stable cutting at *n=3700rpm LentR=2.5*.

5. Entropy

When analysing complex systems with unpredictable behaviour, it is useful to apply the multiscale entropy method. The application of this method improves the understanding of complex phenomena and such analysis is becoming more and more popular [5, 6]. Multiscale entropy is used for determining the complexity of measured finite-length time-series signals.

MSE may be encumbered with some error, depending on the scale factor length τ. Therefore, in this paper composite multi-scale entropy *(CMSE)* is employed, which helps prevent the above-mentioned errors. The concept of composite multiscale entropy is based on the coarse-graining procedure that uses a coarse-grained time series as an average of the original data points within non-overlapping windows by increasing the scale factor τ according to Eq. (6):

$$y_{k,j}^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+\mathrm{k}}^{i=j\tau+k-1} x_i, \quad 1 \le j \le N/\tau, \quad 1 \le k \le \tau. \tag{6}$$

where $\tau=1, 2, 3$, when $\tau=1$, the vector $y_j=x_i$. The vector x is a row of one-dimensional time series. Figure 9 illustrates the averaged values. In this approach, the *CMSE* for each scale factor τ is calculated on the coarse-grained time series $y_{k,j}^{(\tau)}$:

$$CSME(\mathrm{x},\tau,\mathrm{m},\mathrm{r}) = \frac{1}{\tau}\sum_{k=1}^{\tau} SampEn(\mathrm{y}_k^{(\tau)},\mathrm{m},\mathrm{r}) \ . \tag{7}$$

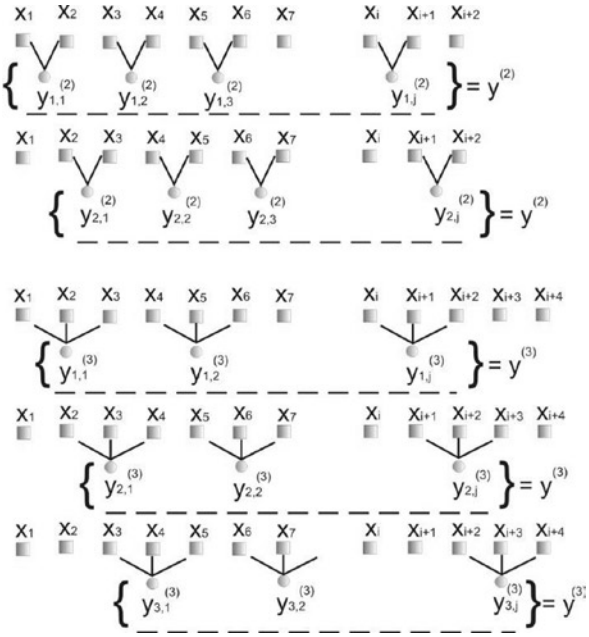


Fig. 9. Diagram of averaged values at the scale index τ=2 and τ=3 in CMSE

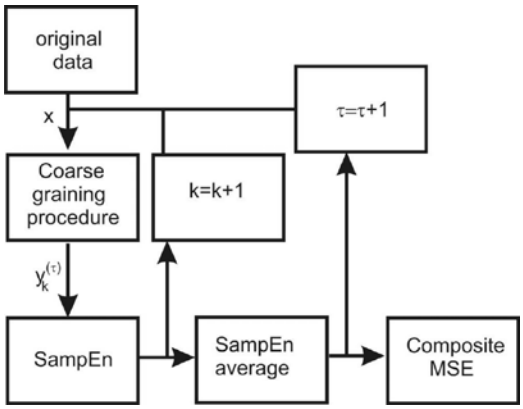


Fig. 10. Algorithm of composite multi-scale entropy (CMSE) calculation

When estimating SampEn in Eq. 7, the values of N_d and N_n should be calculated from prepared grained data $y(\tau)$ by the procedure defined in Eq. 8.

$$\begin{aligned} &N_d = N_n = 1 \\ &if \quad \left|y^{(\tau)}(i) - y^{(\tau)}(j)\right| < r \;\; \& \;\; \left|y^{(\tau)}(i+1) - y^{(\tau)}(j+1)\right| < r \\ &\qquad N_n = N_n + 1; \\ &if \quad \left|y^{(\tau)}(i+2) - y^{(\tau)}(j+2)\right| < r \\ &\qquad N_d = N_d + 1; \end{aligned} \tag{8}$$

Finally, *SampEn* is the logarithm of conditional probability that two sequences with a tolerance r from the points that remain within r to each other at the next point:

$$SampEn(y_k^{(\tau)}, m, r) = \ln\left(\frac{N_n}{N_d}\right). \tag{9}$$

In Eq. 9, the parameter m denotes the length of two likelihood occurrence chains that are similar toward each other in the tolerance r. For the analysis of the time courses $m=2$ is applied, whereas the tolerance of probability $r=0.1\sigma_x$ is applied where σ_x is a standard deviation of the original time course vector. Finally, the *CMSE* is calculated according to the algorithm shown in Fig.10.

The *CMSE* signal analysis was applied to the measured signals of the cutting force in *x*-direction for different sets of two cutting process parameters: the cutting depth a_p and the angular velocity of the spindle n. Figs. 11 and 12 show the *CMSE* results. The composite multiscale entropy analysis revealed that the force signals demonstrate different levels of regularity. The case of machining with the angular velocity of spindle set to $n=4150rpm$ and the cutting depth $a_p=0.2mm$ produced the most regular signals. In other cases, the *CMSE* reaches higher values, which points to irregularity of the measured signals. With the cutting depth set to $a_p=0.05mm$, vibration occurs in the process, irrespective of the spindle velocity n. In Figure 11a one can observe a similar *CMSE* for each scale factor τ, while Fig. 11b reveals significant differences for milling at $n=3000rpm$, $n=3700rpm$ or $n=4150rpm$.

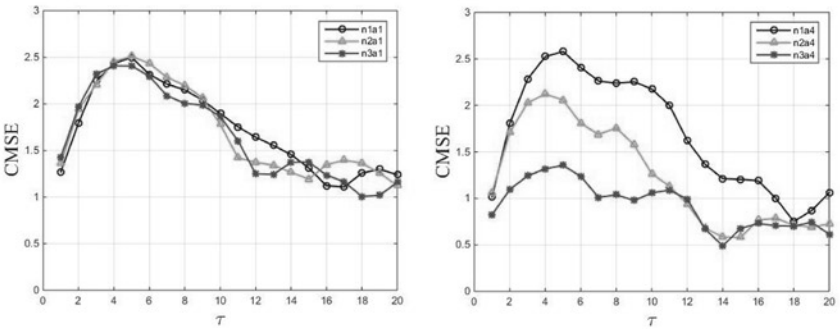


Fig. 11. Composite multiscale entropy analysis calculated for F_x force signals for two cutting depths a1: a_p=0.05mm (a) and a4: a_p=0.20mm (b) compared to spindle rotational speeds n1: n=3000rpm, n2: n=3700rpm and n3: n=4150rpm.

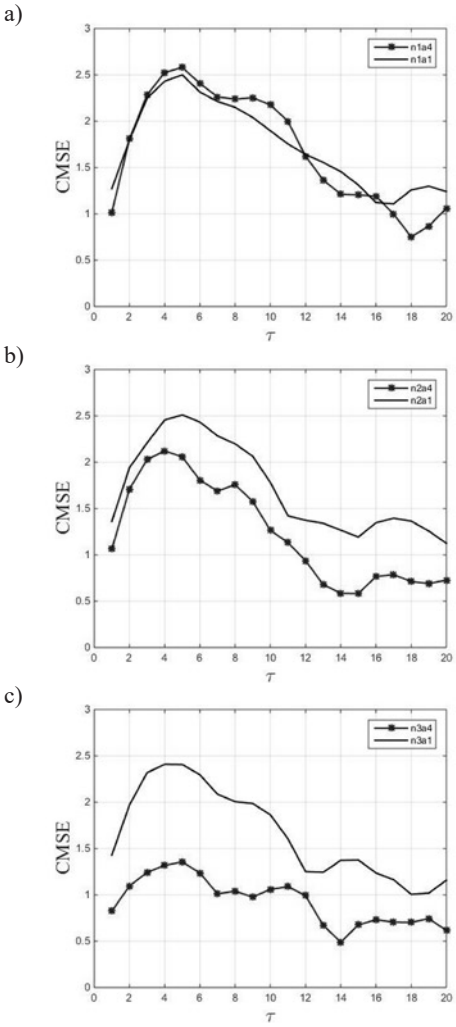


Fig. 12. Composite multiscale entropy analysis calculated for F_x force signals for different ranges of the cutting depth (a1: a_p=0.05mm and a4: a_p=0.20mm) and rotational speeds of spindle n1: n=3000rpm (a), n2: n=3700rpm (b) and n3: n=4150rpm (c)

The above is clearly shown in Figs. 12a, b and Fig. 12c, where it is plotted with respect to an increased angular velocity. Consequently, the signals are compared between the cases of inside and outside lobes shown in Fig. 2b. This corresponds to the stable and unstable conditions of the milling process. One can notice that the unstable case shown in Fig. 2 (point n3a4) turns out to be a more regular condition for machining (stars line in Fig. 11b or circles line in Fig.12c) than the second set point n3a1 (stars line in Fig. 11a or black line in Fig. 12c). Moreover, when the angular velocity of the spindle is decreased to $n=3700rpm$ (Fig. 12b) or $n=3000rpm$ (Fig.12a), the irregularity of the signals is similar to the case n3a1 ($n=4150rpm$, $a_p=0.05mm$ in Fig.12c). The most significant difference can be observed for the highest analysed angular velocity (n3). The system is sensitive to process parameters and the *CMSE* shows a clear discrepancy whether the process is stable or not. For the angular velocity $n=3700rpm$ (Fig. 12b) or even $n=3000rpm$ (Fig.12a), the *CMSE*

is similar, which indicates that the depth of cutting has no significant effect on machining stability. Finally, one can observe that the *CMSE* method provides an alternative approach to estimating conditions of the machining process.

6. Conclusion

The paper investigated the problem of the milling process stability for nickel alloy. First, a modal analysis was performed to obtain a stability lobes diagram; the diagram was then verified for selected points by the statistical method, recurrence technique and multiscale entropy method.

Based on the statistical analysis results, it is found that in the case of unstable cutting (points above the stability lobes), there is a significant increase in the maximum cutting force, particularly in the feed direction *(x)*. Similar conclusions can be reached by analysing the standard deviation that is quite obvious in this case. However, it is interesting to observe that a greater symmetry and concentration of the analysed signals occur in the unstable points. Recurrence plots and especially *RQA* are useful for identification chatter vibrations in cutting processes. Given a variety of *RQA* indexes, only a few of them were tested here to select the best one. Undoubtedly, the Shannons entropy (Lent) is the best because it can distinguish stable and unstable cutting for various rotational speeds (cutting speeds). The Lent ratio (between unstable and stable cutting) is the best indicator of cutting instability. *L*-entropy is extremely high when chatter vibrations appear (unstable cutting), because chatter vibrations show a higher regularity than small-amplitude stochastic vibrations in stable cutting. The recurrence diagrams are not useful to a sufficient degree because it is difficult to estimate differences in the *RPs* and the differences are subjective.

Independently of the above, the composite multiscale entropy analysis revealed a certain degree of signal regularity, which may prove useful when estimating which can give a hint if the milling conditions are profitable for the process. As regards the analysed force signals, it was noticed that increasing the rotational speed of the spindle causes the system to behave in a more regular way, provided that the cutting depth has been selected properly. This observation can be useful when other methods such as *RP* and *RQA* fail or their results are questionable. Since some of the analysed indexes show quite significant differences between stable and unstable cutting, the proposed methods have potential to be employed in the future in a chatter control system.

Acknowledgement

The work of the first author was financially supported by the National Science Centre under the project no. DEC- 2013/09/N/ST8/01202. The second and third authors' research was supported by the Polish Ministry of Science and Higher Education (S3/M/2016). The contribution of the fourth author was financially supported by the National Science Centre under the project no. DEC-2011/01/B/ST8/07504.

References

1. Adetoro O, Sim W, Wen P. An improved prediction of stability lobes using nonlinear thin wall dynamics. Journal of Materials Processing Technology 2010; 210: 969-979, https://doi.org/10.1016/j.jmatprotec.2010.02.009.
2. Altintas Y, Budak E. Analytical prediction of stability lobes in milling. CIRP Annals - Manufacturing Technology 1995; 44: 357-362, https://doi.org/10.1016/S0007-8506(07)62342-7.
3. Altintas Y, Budak E. Analytical prediction of chatter stability in milling-part I: general formulation. Journal of Dynamic Systems, Measurement, and Control 1998; 120: 22-30, https://doi.org/10.1115/1.2801317.
4. Bayly P V, Halley J E, Mann B P, Davies M A. Stability of interrupted cutting by temporal finite element analysis. Journal of Manufacturing Science and Engineering 2003; 125: 220-225, https://doi.org/10.1115/1.1556860.
5. Borowiec M, Sen A K, Litak G, Hunicz J, Koszałka G, Niewczas A. Vibrations of a vehicle excited by real road profiles. Forschung im Ingenieurwesen 2010; 74(2): 99-109, https://doi.org/10.1007/s10010-010-0119-y.
6. Borowiec M, Rysak A, Betts D, Bowen C, Kim H, Litak G. Complex response of a bistable laminated plate: Multiscale entropy analysis. European Physical Journal Plus 2014; 129(211): 1-7, https://doi.org/10.1140/epjp/i2014-14211-3.
7. Davies M A, Balachandran B. Impact dynamics in milling of thin-walled structures. Nonlinear Dynamics 2000; 22: 375-392, https://doi.org/10.1023/A:1008364405411.
8. Ding Y, Zhu L, Zhang X, Ding H. A full-discretization method for prediction of milling stability. International Journal of Machine Tools and Manufacture 2010; 50: 502-509, https://doi.org/10.1016/j.ijmachtools.2010.01.003.
9. Eckmann J P, Kamphorst S O, Ruelle D. Recurrence plots of dynamical systems. Europhysics Letters (EPL) 1987; 4: 973-977, https://doi.org/10.1209/0295-5075/4/9/004.
10. Gang L. Study on deformation of titanium thin-walled part in milling process. Journal of Materials Processing Technology 2009; 209: 2788-2793, https://doi.org/10.1016/j.jmatprotec.2008.06.029.
11. Gradisek J, Govekar E, Grabec I. Chatter onset in non-regenerative cutting: a numerical study. Journal of Sound and Vibration 2001; 242: 829-838, https://doi.org/10.1006/jsvi.2000.3388.
12. Insperger T. Act-and-wait concept for time-continuous control systems with feedback delay. IEEE Transactions on Control Systems Technology 2006; 14: 974-977, https://doi.org/10.1109/TCST.2006.876938.
13. Izamshah R, Mo J, Ding S. Hybrid deflection prediction on machining thin-wall monolithic aerospace component. Proceedings of the Institution of Mechanical Engineers Part B-Journal of Engineering Manufacture 2012; 226: 592-605, https://doi.org/10.1177/0954405411425443.
14. Kęcik K, Rusinek R, Warmiński J. Stability lobes analysis of nickel superalloys milling. International Journal of Bifurcation and Chaos 2011; 21: 1-12, https://doi.org/10.1142/S0218127411030258.
15. Kęcik K, Rusinek R, Warmiński J, Weremczuk A. Chatter control in the milling process of composite materials. Journal of Physics: Conference Series 2012; 382: 012012, https://doi.org/10.1088/1742-6596/382/1/012012.
16. Kęcik K, Borowiec M, Rusinek R. Verification of the stability lobes of Inconel 718 milling by recurrence plot applications and composite multiscale entropy analysis. The European Physical Journal Plus 2016; 131: 27-36.
17. Krauze K, Jankowski Z, Blaschke J. Cutting forces eyaluation for tangentiar-rotational tool basing on laboratory experiments. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2004; 24: 9-14.
18. Litak G, Syta A, Rusinek R. Dynamical changes during composite milling: recurrence and multiscale entropy analysis. International Journal of Advanced Manufacturing Technology 2011; 56: 445-453, https://doi.org/10.1007/s00170-011-3195-8.
19. Litak G, Rusinek R. Dynamics of a stainless steel turning process by statistical and recurrence analyses. Meccanica 2012; 47: 1517-1526,

https://doi.org/10.1007/s11012-011-9534-x.
20. Litak G, Kęcik K, Rusinek R. Cutting force response in milling of Inconel: Analysis by wavelet and Hilbert-Huang transforms. Latin American Journal of Solids and Structures 2013; 10: 133-140, https://doi.org/10.1590/S1679-78252013000100013.
21. Litak G, Polyakov Y S, Timashev S F, Rusinek R. Dynamics of stainless steel turning: Analysis by flicker-noise spectroscopy. Physica A: Statistical Mechanics and its Applications 2013; 392: 6052-6063, https://doi.org/10.1016/j.physa.2013.07.079.
22. Lopez Lacalle L N. Machine tools for high performance machining. London: Springer, 2009, https://doi.org/10.1007/978-1-84800-380-4.
23. Mane I, Gagnol V, Bouzgarrou B, Ray P. Stability-based spindle speed control during flexible workpiece high-speed milling. International Journal of Machine Tools and Manufacture 2008; 48: 184-194, https://doi.org/10.1016/j.ijmachtools.2007.08.018.
24. Mann B P, Insperger T, Bayly P V, Stepan G. Stability of up-milling and down-milling, part 2: experimental verification. International Journal of Machine Tools and Manufacture 2003; 43: 35-40, https://doi.org/10.1016/S0890-6955(02)00160-8.
25. Marwan N, Kurths J. Nonlinear analysis of bivariate data with cross recurrence plots. Physics Letters A 2002; 302: 299-307, https://doi.org/10.1016/S0375-9601(02)01170-2.
26. Marwan N, Thiel M, Nowaczyk N R. Cross recurrence plot based synchronization of time series. Nonlinear Processes in Geophysics 2002; 9: 325-331, https://doi.org/10.5194/npg-9-325-2002.
27. Merdol S D, Altintas Y. Multi frequency solution of chatter stability for low immersion milling. Journal of Manufacturing Science and Engineering-Transactions of the Asme 2004; 126: 459-466, https://doi.org/10.1115/1.1765139.
28. Peigne G, Paris H, Brissaud D, Gouskov A. Impact of the cutting dynamics of small radial immersion milling operations on machined surface roughness. International Journal of Machine Tools and Manufacture 2004; 44: 1133-1142, https://doi.org/10.1016/j.ijmachtools.2004.04.012.
29. Rusinek R. Vibrations in cutting process of titanium alloy. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2010; 3: 48-55.
30. Rusinek R. Cutting process of composite materials: An experimental study. International Journal of Non-Linear Mechanics 2010; 45: 458-462, https://doi.org/10.1016/j.ijnonlinmec.2010.01.004.
31. Rusinek R, Borowiec M. Stability analysis of titanium alloy milling by multiscale entropy and Hurst exponent. The European Physical Journal Plus 2015; 130: 194, https://doi.org/10.1140/epjp/i2015-15194-1.
32. Rusinek R, Lajmert P, Kęcik K, Kruszyński B, Warmiński J. Chatter identification methods on the basis of time series measured during titanium superalloy milling. International Journal of Mechanical Sciences 2015; 99: 196-207, https://doi.org/10.1016/j.ijmecsci.2015.05.013.
33. Rusinek R, Zaleski K. Dynamics of thin-walled element milling expressed by recurrence analysis. Meccanica 2016; 51: 1275-1286, https://doi.org/10.1007/s11012-015-0293-y.
34. Seguy S, Dessein G, Arnaud L. Surface roughness variation of thin wall milling, related to modal interactions. International Journal of Machine Tools and Manufacture 2008; 48: 261-274, https://doi.org/10.1016/j.ijmachtools.2007.09.005.
35. Semotiuk L. An analysis of the operational characteristics of innovative tool structures used in high speed rough milling processes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2009; 41: 46-53.
36. Schmitz T L, Medicus K, Dutterer B. Exploring once-per-revolution audio signal variance as a chatter indicator. Machining Science and Technology 2002; 6: 215-233, https://doi.org/10.1081/MST-120005957.
37. Thevenot V, Arnaud L, Dessein G, Cazenave-Larroche G. Integration of dynamic behaviour variations in the stability lobes method: 3D lobes construction and application to thin-walled structure milling. The International Journal of Advanced Manufacturing Technology 2006; 27:638-644, https://doi.org/10.1007/s00170-004-2241-1.
38. Voronov S, Kiselev I. Dynamics of flexible detail milling. Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics 2011; 225: 299-309, https://doi.org/10.1177/1464419311418735.
39. Wiercigroch M, Budak E. Sources of nonlinearities, chatter generation and suppression in metal cutting. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 2001; 359: 663, https://doi.org/10.1098/rsta.2000.0750.
40. Wiercigroch M, Krivtsov A M. Frictional chatter in orthogonal metal cutting. Phil.Trans.The Royal Society Society of London A Mathematical Physical and Engineering Science 2001; 359: 713-738.
41. Wojciechowski S, Twardowski P, Pelic M, Maruda R W, Barrans S, Krolczyk G M. Precision surface characterization for finish cylindrical milling with dynamic tool displacements model. Precision Engineering 2016; 46: 158-165, https://doi.org/10.1016/j.precisioneng.2016.04.010.
42. Wojciechowski S, Maruda R W, Barrans S, Nieslony P, Krolczyk G M. Optimisation of machining parameters during ball end milling of hardened steel with various surface inclinations. Measurement 2017; 111: 18-28, https://doi.org/10.1016/j.measurement.2017.07.020.
43. Zębala W, Słodki B, Struzikiewicz G. Productivity and reliability improvement in turning Inconel 718 alloy - case study. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15: 421-426.
44. Zbilut J P, Webber C L J. Embeddings and delays as derivedfrom quantification of recurrence plots. Physics Letters A 1992; 171: 199-203, https://doi.org/10.1016/0375-9601(92)90426-M.
45. Zhenyu S, Luning L, Zhanqiang L. Influence of dynamic effects on surface roughness for face milling process. The International Journal of Advanced Manufacturing Technology 2015; 80: 1823-1831, https://doi.org/10.1007/s00170-015-7127-x.

Andrzej WEREMCZUK
Marek BOROWIEC
Michał RUDZIK
Rafał RUSINEK
Department of Applied Mechanics
Lublin University of Technology
ul. Nadbystrzycka 36, 20-816 Lublin, Poland

E-mail: a.weremczuk@pollub.pl, m.borowiec@pollb.pl, m.rudzik@pollub.pl, r.rusinek@pollub.pl

Article citation info:
BABISHIN V, HAJIPOUR Y, TAGHIPOUR S. Optimisation of non-periodic inspection and maintenance for multicomponent systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (2): 327–342, http://dx.doi.org/10.17531/ein.2018.2.20.

Vladimir BABISHIN
Yassin HAJIPOUR
Sharareh TAGHIPOUR

OPTIMISATION OF NON-PERIODIC INSPECTION AND MAINTENANCE FOR MULTICOMPONENT SYSTEMS

OPTYMALIZACJA NIE-OKRESOWYCH PRZEGLĄDÓW I KONSERWACJI SYSTEMÓW WIELOELEMENTOWYCH

A k-out-of-n:G system and a system with components subject to soft and hard failures are both inspected non-periodically. For the k-out-of-n system, components fail "silently" (i.e. are hidden), and the entire system fails when (n-k+1)st component fails. For the system with hard-type and soft-type components, hard failures cause system failure, while soft failures are hidden and do not cause immediate failure of the system, but still reduce its reliability. Every system failure allows for an opportunistic inspection of hidden soft-type components in addition to the scheduled inspections. The available maintenance types are replacement and minimal repair. For hard-type components, the maintenance decision is determined by the optimal age before replacement. For the soft-type components with hidden failures, we do not know their age, and so decide on the appropriate type of maintenance using the optimal number of minimal repairs before replacement. The hidden nature of soft-type component failures precludes the use of a tractable analytic expression, so we use simulation and genetic algorithm (GA) to jointly optimise the non-periodic policies on maintenance and inspection and to ensure these incur minimal expected total cost over a finite planning horizon. Due to increasing computational complexity associated with the number of inspections and maintenance policies to be evaluated, the genetic algorithm presents a promising method of optimisation for complex multicomponent systems with multiple decision parameters.

Keywords: *non-periodic inspection, opportunistic inspection, maintenance, hidden soft failure, hard failure, genetic algorithm.*

Przeglądów systemu typu k z n: G oraz systemu z elementami ulegającymi miękkim i twardym uszkodzeniom dokonuje się nieokresowo. W przypadku systemu k z n, składowe ulegają uszkodzeniom „w trybie cichym" (tj. uszkodzenia są ukryte), a cały system ulega awarii, gdy ulegnie uszkodzeniu element (n-k + 1). W przypadku systemu z elementami typu twardego i miękkiego, uszkodzenia twarde prowadzą do awarii systemu, natomiast uszkodzenia miękkie są ukryte i nie powodują natychmiastowej awarii systemu, choć nadal zmniejszają jego niezawodność. Każda awaria systemu stanowi dodatkową, w stosunku do przeglądów planowych, okazję do przeprowadzenia przeglądu (tzw. przegląd awaryjny) ukrytych elementów miękkich. Dostępne rodzaje konserwacji to wymiana oraz naprawa minimalna. W przypadku komponentów twardych, decyzję, który typ konserwacji zastosować, podejmuje się biorąc pod uwagę optymalny wiek przed wymianą. W przypadku elementów miękkich z ukrytymi uszkodzeniami, wiek optymalny jest nieznany, dlatego decyzje o odpowiednim typie konserwacji podejmuje się z uwzględnieniem optymalnej liczby minimalnych napraw przed wymianą. Ukryty charakter uszkodzeń elementów składowych typu miękkiego wyklucza wykorzystanie rozwiązywalnego wyrażenia analitycznego, dlatego w pracy użyto symulacji i algorytmu genetycznego (GA), w celu jednoczesnej optymalizacji nieokresowych strategii prowadzenia konserwacji i przeglądów oraz zapewnienia, że będą one pociągały za sobą minimalny oczekiwany koszt całkowity w skończonym horyzoncie planowania. W świetle rosnącej złożoności obliczeniowej związanej z dużą liczbą ocenianych przeglądów i strategii utrzymania ruchu, algorytm genetyczny stanowi obiecującą metodę optymalizacji złożonych systemów wieloelementowych o wielu parametrach decyzyjnych.

Słowa kluczowe: *przegląd nie-okresowy, przegląd awaryjny, utrzymanie ruchu, ukryte uszkodzenie miękkie, uszkodzenie twarde, algorytm genetyczny.*

Notation

Typesetting Convention: vectors, matrices and arrays are bold.

Latin Symbols

C	Cost (random variable).
E	Expectation.
F	Expected number of system failures.
G	Generator function for the expected values.
M	Expected number of minimal repairs.
P	Probability.
R	Expected number of replacements.
U	Expected uptime.
UCL	Upper confidence limit.
Y	First failure time for a soft-type component.
Z	First failure time for hard-type component subsystem.
a	Inspection policy index.
b	Random number.
c	Cost (constant).
f	Number of component failures.
h	Hard-type component index.
i	Scheduled inspection index.
j	Component index.

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

k Number of identical components required to be in working order for a k-out-of-n system.
l Total number of inspections within a planning horizon.
m Number of minimal repairs before replacement.
n Total number of components.
s Soft-type component index.
t Age of a component.
x Inspection policy.
$\mathcal{R}$ Reliability.

Greek Symbols

Φ Random variable for the number of component failures.
Λ Cumulative hazard function.
Y Uptime of a component.
α Confidence level.
β Shape parameter in Weibull distribution.
ζ Optimal replacement age for hard-type component.
η Scale parameter in Weibull distribution, characteristic life.
θ Age of a hard-type component.
λ Power-law intensity (hazard) function.
σ Time until next inspection.
τ Minimal time unit.
φ Instance number of component failures.
χ Time-to-failure.
ω Length of planning horizon (life cycle).

Superscripts

* Optimality.
D Component downtime.
I Inspection.
M Minimal repair.
R Replacement (see "corrective replacement").
T Total.
HS System composed of hard-type and soft-type components.
SD System downtime.
a Inspection policy identifier.
k,n System composed of k-out-of-n components.

Subscripts

1 Reference to soft-type components.
2 Reference to hard-type components.
c Current.
h Hard-type component identifier.
i Scheduled inspection identifier.
j Component identifier.
l Reference to the total number of inspections within a planning horizon.
m Number of minimal repairs before replacement maintenance criterion.
s Soft-type component identifier.

1. Introduction and Background

Multicomponent systems generally have higher complexity than unicomponent systems, since the former usually have one or more intercomponent dependencies, such as functional, structural, failure, or economic [8, 28, 34]. Optimal maintenance and economic dependency in multicomponent systems is studied by Dekker et al. [11], Wang and Pham [36] and Zille et al. [41]. Periodic replacement policies for multicomponent systems with stochastic and economic dependencies are investigated by Ozekici [20]. Series systems with mixed standby components are compared based on their cost/benefit ratio, time to failure and long-term availability by Wang and Kuo [35].

Redundant systems with high levels of availability, reliability and robustness are typically configured as *k*-out-of-*n* systems, where the system is able to perform without interruption until failures of its components accumulate to *n*-*k*+1. Multi-engine aircraft, multi-display airplane cockpits, dual-contour automotive brake lines and multiple pumps used for hydraulic control are just several examples of *k*-out-of-*n* systems. A *k*-out-of-*n* system with perfect component repairs and maintenance equipment subject to imperfect repairs is considered by Zhang and Wu [38]. Load-sharing *k*-out-of-*n* systems are considered by Taghipour [27] and Taghipour and L. Kassaei [32]. They minimise the total expected cost and determine the optimal inspection interval for a finite planning horizon.

Definition 1.1: Generally, failure is an adverse event, which interferes with the normal designed functioning of the affected unit. One major class of multicomponent systems includes those composed of the two types of components classified by failure: hard-type and soft-type.

Definition 1.2: A *hard-type component* is a component whose failure is self-evident and triggers the system failure immediately; therefore, the time of failure is known for this component type. Examples of hard-type components include: wiring in ignition distributor in automotive electronic ignition, central processing unit in personal computers, fuse and display in infusion pumps, etc.

Definition 1.3: A *soft-type component* is a component whose failure does not trigger the immediate system failure, but the latter's reliability is usually reduced as a result of increased risk of malfunction, damage and/or eliminated redundancy. We refer to failures proper as "hard and soft failures" whenever the focus is on the failure process, and to components proper as "hard-type and soft-type components", respectively, to distinguish between different types and behaviours of components.

Definition 1.4: *System reliability* means the probability that the system will operate without failure under the design operating conditions (such as voltage, temperature, humidity). *Component reliability* refers to the same concept applied to individual components, whether hard-type, or soft-type.Examples of soft-type components include: liquid-level alarms in infusion pumps and standby-redundant components (batteries, surge-protective equipment, parallel processors).

Parts of the system subject to both soft and hard failures are treated as separate components of different types. It can be also noted that components in *k*-out-of-*n* systems can be treated as soft-type due to the system's capacity to accumulate component failures.

Periodic inspection policies for complex multicomponent systems have been extensively studied by Taghipour and Banjevic [31, 30, 29], Flage and Aven [12] and Pandey et al. [21]. Taghipour and Kassaei [32] consider periodic inspection optimisation for *k-out-of-n* systems.

For almost any system, the planning horizon is related to the system's life expectancy, depending on the operational and/or managerial objectives. Fixed and finite planning horizon is used in areas such as pharmacology, medical devices with expiry date, aircraft maintenance (Sriram and Haghani [25]). For example, medications and a vast majority of medical tools have to be replaced once the end of their life cycle has passed. Similarly, aircraft parts usually have to be preventively replaced after a specific number of flight hours.

Systems such as protective devices usually contain components whose failures are hidden. A hidden failure is a failure revealed only at inspection, but not during the normal operation of the system [19]. The detection of a hidden failure in an integrated system composed of main functional (protected) and safety (protective) units may occur either at inspection, or whenever the protective unit is required to function, but is unavailable because of a failure. Soft failures are similar to hidden failures, but the system is still able to function despite their presence. Single-component systems with hidden failures, probability of failure dependent on the number of previous repairs,

and maintenance policy based on both the component's age-at-failure and its number of overhauls are investigated by Sheu et al. [24].

Bjarnason et al. [6] consider a joint optimisation model for minimising the total cost of both maintenance and inventory policies for a system with hidden failures in a *k*-out-of-*n* redundant configuration. Babishin and Taghipour [2] propose a joint optimisation procedure for minimising the total cost of periodic inspections and additional inspections at system failures (opportunistic inspections), as well as corrective maintenance in *k*-out-of-*n* systems, where component failures are hidden.

Failure of an entire system or some of its components can be regarded as an opportunity to check all of the components for damage in addition to the scheduled inspections – hence, whenever such opportunity is taken, inspections performed at that time are called "opportunistic". In the literature, opportunistic maintenance has received an extensive treatment. For example, Dagpunar [10] considers a prespecified control limit for component's age, exceeding which a failed component in a multicomponent system is opportunistically replaced. Zhu et al. [40] offer a policy for opportunistic maintenance of offshore wind turbines with two component types, where maintenance action for soft-type components depends on their ages. Cui and Li [9] model damage in a multicomponent system accumulating to a shock event under opportunistic inspections and stochastically-dependent components. Aven and Dekker [1] consider age-based, as well as block replacement models with opportunities for preventive replacement. Gao et al. [13] propose a quasi-periodic imperfect preventive maintenance policy for a repairable system with stochastic maintenance interval. Peng et al. [23] study a sequential periodic preventive maintenance policy and develop a hybrid random imperfect maintenance model, optimising it using genetic algorithm. Legát et al. [18] consider both periodic preventive and predictive maintenance and determine, correspondingly, the optimal interval and the optimal diagnostic parameter. Gunn and Diallo [15] use a shortest path depth-first algorithm to search a network tree representation of the indirect opportunistic grouping of preventive periodic replacements. Yun and Endharta [37] use minimal cut set to analyse a *k*-out-of-*n*:F system with exponential failure times and evident failures.Unlike the cases from the literature, in the present case, there is a choice of maintenance action which the maintenance personnel may take at every failure of a component.

Genetic algorithms (GA) have been used in the literature for inspection optimisation of multicomponent systems. Because of the absence of analytical solution, Babishin and Taghipour [3] employ joint optimisation with exhaustive search, as well as genetic algorithm with either integer, or quasi-continuous inspection period. They provide the optimal joint inspection and maintenance policies, as well as calculate the expected number for system failures depending on the cost ratio and hazard function of components in a k-out-of-n system with hidden component failures under preventive or corrective replacement, minimal repair, opportunistic and periodic inspections. They also derive a criterion for calculating the acceptable number of system failures over its planning horizon. Bjarnason and Taghipour [5] formulate a model for a *k*-out-of-*n* system with a three-dimensional objective function and use the GA to find the joint optimal (s, S) inventory and maintenance policies. It should be noted that analysing systems with mixed hard and soft failures, where each component may require a special treatment, is generally much harder than analysing systems in *k*-out-of-*n* configuration with identical components.

Non-periodic inspection optimisation has also been covered in the literature. Zhao et al. [39] consider a system made of one component, which is replaced under a Gamma deterioration process over infinite planning horizon and develop a proportional hazards model for optimising the system's non-periodic inspection policy. Su [26] essentially develops a model for optimising non-periodic inspections in a one-component system with a combination of hidden and self-announcing operating modes, since his inspection "period" is a random variable, which renders it non-periodic according to the definitions and terminology adopted in the present paper. He uses the supplementary variable technique to optimise for an inspection period which maximises profit per unit time.

Multicomponent non-periodic inspections have also been considered in the literature. Hajipour and Taghipour's model [16] optimises for non-periodic inspection policy in a finite life cycle for multicomponent systems with a choice from two maintenance actions performed based on the age-dependent probability. Castanier et al. [7] propose a model taking into account the condition of the system for optimal inspection and replacement of a two-component system under non-periodic inspections, where they essentially develop separate policies for each component, assuming component independence, admitting that extending their approach to larger systems makes the numerical solution intractable. In this regard, it is worth mentioning that Vaurio notes in [33] that it is not generally possible to obtain an analytical solution for the optimal inspection interval even in the simpler case of optimising only for system availability. This explains the interest in and the value of numerical and simulation methods for the analysis of multicomponent systems.

Golmakani and Moakedi [14] develop a model for non-periodic inspection optimisation using branch-and-bound and dynamic programming techniques, which they use to introduce the A* search algorithm, which attempts to improve on the efficiency of branch-and-bound technique using branching on the most attractive nodes at each step in the procedure. However, the A*search is at a disadvantage for generating a large number of nodes at each iteration. Some researchers, e.g. Lapa et al. [17], demonstrated the applicability and usefulness of genetic algorithms to optimisation of system availability. In the present paper, genetic algorithm is used for the purpose of improving efficiency of optimisation calculations.

In summary, the present paper provides a general methodology and two models for finding the optimal joint non-periodic inspection and maintenance policies for complex multicomponent systems with finite planning horizon. In the previous models such as, for example, by Hajipour and Taghipour [16], Taghipour and Banjevic [31, 30, 29], the maintenance action was not optimised, and failed components were replaced, or minimally repaired based on age-dependent probability. Babishin and Taghipour [3]optimise both maintenance and inspection policies for a system in *k*-out-of-*n* configuration, but only under periodic inspections. Babishin and Taghipour [4] use a three-stage optimisation procedure to obtain optimal inspection policy for hard-type components in Stage 1, optimal maintenance in Stage 2 and optimal periodic inspection interval for soft-type components in Stage 3 using the Monte Carlo simulation.

In the present paper, both the maintenance decision and the inspection policy are optimised jointly in one stage. Recursive mathematical formulations for generating the expected values of minimal repairs, replacements and uptime are also provided for the first time in the case of a *k*-out-of-*n* system.The optimal maintenance policy for soft-type components is determined by the number of minimal repairs until replacement for these components, similarly to the approach proposed by Park [22]. The optimal maintenance policy for components with hard failures is based on these components' ages. Both of the proposed models feature corrective maintenance (replacement or minimal repair) of components with hard and soft failures, along with scheduled non-periodic and opportunistic inspections of components with soft failures. The hard failure occurrence in the system composed of hard- and soft-type components affects the expected number of soft failures, replacements, minimal repairs and expected downtime. Therefore, these expected values influence the optimal inspection policy. The components of a *k*-out-of-*n* system are regarded as being identical soft-type components, which facilitates the analysis of such systems. Jointly optimising for both inspection and maintenance in one stage for both systems allows finding optimal maintenance and

inspection policies for entire systems rather than marginally just for certain groups of components.

Generally, the safe and reliable operation of different equipment can be facilitated with the help of inspection and maintenance optimisation models. The latter also have strong managerial implications due to the importance of justifying these decisions with both qualitative and quantitative analysis. Using the proposed inspection and maintenance optimisation models, the decision-maker(s) gain an opportunity to find the combination of inspection and maintenance decisions that is most likely to result in the greatest cost savings without sacrificing availability or reliability. The hard-to-quantify effects, such as those of opportunistic inspections, can be accounted for by using the joint optimisation models in managerial decision-making process. This is likely to result in cost savings, which are especially significant, if the costs of inspection are high. Thus, it can be seen that optimisation of inspection and maintenance decisions represents a valuable asset for decision-makers.

The present article is further organised as following: Section 2 states the problem description; Section 3 outlines the model formulation for systems in *k*-out-of-*n* configurations under non-periodic and opportunistic inspections; Section 4 contains the model formulation for the system composed of hard-type and soft-type components under non-periodic inspections and opportunistic inspections of soft-type components; Section 5 illustrates the models by providing numerical examples; lastly, Section 6 summarises the conclusions.

2. Problem Description

Consider the problem of inspecting devices consisting of coupled systems, such as surge-protected personal computers (PCs), infusion pumps with liquid-level alarms, generators or power distributors with reserve power supplies. For such systems, it may not be economically feasible to have periodic inspections – for example, in the case when the optimal inspection period of the protective system does not coincide with the inspection period of the system they are coupled to. In such cases, non-periodic inspections are a good option.

In the present article, two main kinds of multicomponent systems are considered, based on the classification by the types of component failures. The system belonging to the first kind (System 1) is a *k*-out-of-*n* system with hidden component failures identifiable solely at inspections. The system belonging to the second kind (System 2) consists of components belonging to either of the two types: hard type, or hidden soft type. Both kinds of systems are considered in more detail in subsequent sections.

The present paper focuses on finding the optimal non-periodic policies for maintenance and inspection of two kinds of multicomponent systems described above. The relevant assumptions pertinent to System 1 are further identified by designation "S.1.#",those pertinent solely to System 2 – by "S.2.#", and those pertinent to both systems – by "S.1/2.#", where "S." stands for "system"). We start with stating the general assumptions for both kinds of systems:

S.1/2.1: Soft failures are discovered only at inspections. Therefore, the ages at failure of soft-type components are unknown.

S.1/2.2: Inspections are considered as being either scheduled non-periodic, or opportunistic.

S.1/2.3: Systems are always inspected at the end of the fixed planning horizon, all necessary maintenance is performed and components' ages recorded at that time in order to create a renewal point, after which the optimisation procedure can be repeated again.

S.1/2.4: Scheduled non-periodic (further referred to as simply "non-periodic") inspections occur with the minimal unit of time over a finite planning horizon ω, possibly at times $i\tau$, i=1,2,…,l, $l\in\mathbb{N}$, where $l=\omega/\tau$-1 if ω is divisible by τ, and $l=|\omega/\tau|$ otherwise. Scheduled inspections are always performed on the operating (unfailed) system at times prescribed by an inspection policy.

S.1/2.5: System failures offer an occasion to inspect every component in a system. Every failed component is then maintained to restore its functionality.

S.1/2.6: A maintenance action is classified as either a minimal repair, or a corrective replacement (further referred to as simply "replacement"). A minimal repair restores the component's functionality to the state it was in just preceding the component's failure, thus leaving the component's age unaffected. A corrective replacement decreases the failed component's age to 0 ("as-good-as-new" state). Minimal repairs and replacements can take place at scheduled, as well as opportunistic inspections.

S.1/2.7: Both maintenance and inspection are assumed to have negligible duration and are perfect.

When obtaining the total expected cost for a number of minimal repairs before replacement m_j per component within a given planning horizon, we assume the following:

S.1/2.8: Component j is replaced after m_j minimal repairs.

Furthermore, for the purposes of calculating the upper bound for the expected number of component failures within the planning horizon at α confidence level for each soft-type component, we assume that:

S.1/2.9: There is no delay in detecting failures upon inspection.

S.1/2.10: Component failures are rectified by minimal repair. (No replacements are considered here, since we are interested in finding the optimal number of times a failed component may be minimally repaired before it is replaced for the first time).

The number of minimal repairs before replacement depends on the expected number of component failures, since there is a statistical uncertainty associated with the latter. In obtaining the expected number of component failures $E[\Phi(\omega)]$, we observe the following about the component failure process:

S.1/2.11: $\Phi(t)\in\{\mathbb{N}\cup\{0\}\}$, (i.e. the number of failures is a non-negative integer).

S.1/2.12: If $t_1<t_2$, then $\Phi(t_1)\leq\Phi(t_2)$.

S.1/2.13: For $t_1<t_2$, $\Phi(t_2)-\Phi(t_1)$ equals the number of failures which occurred in the interval (t_1,t_2). Without the loss of generality, if t_1=0, t_2=ω, then $\Phi(t_2)-\Phi(t_1)=\varphi$.

Based on S.1/2.11-S.1/2.13, it can be concluded that the component failure process is a counting process. Furthermore:

S.1/2.14: $\Phi(0)$=0, (i.e. no components have failed prior to the beginning of the system's life cycle).

S.1/2.15: Component failures follow independent increments, i.e. the numbers of component failures in disjoint time intervals are mutually independent of each other.

For System 1, the following assumptions are also made:

S.1.1: The number of components for System 1 is denoted as n.

S.1.2: All redundant components in *k*-out-of-*n* configuration are identical.

S.1.3: Components' failures are assumed to occur according to a power law intensity function (hazard function) $\lambda_j(t)=\frac{\beta_j}{\eta_j}\left(\frac{t_j}{\eta_j}\right)^{\beta_j-1}$ following non-homogeneous Poisson process (NHPP), where β_j is the shape parameter and η_j is the scale parameter of the Weibull distribution describing times between failures of component j, t_j is the age of component j, j=1,2,…,n for System 1.

S.1.4: Opportunistic inspections are incurred whenever n-k+1 components fail.

S.1.5: The (n-k+1)th failure presents an opportunity for inspecting the system and rectifying the failed components, which influences the number of replacements, minimal repairs and downtime for the hidden components (see Fig.1).

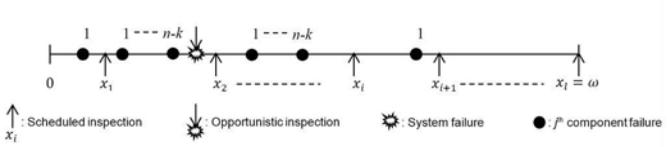


Fig.1: Sample hidden failures, scheduled and opportunistic inspections within one life cycle for System 1.

Consider the problem of optimising inspection and maintenance of several identical antennae providing network access on a remote base station (e.g. used in geophysical surveys), which because of crew staffing shortages, weather, or accessibility issues requires non-periodic inspection and maintenance. Each antenna is considered as a component, and the collection of antennae providing network access are considered as a system. When an antenna's battery runs down, it ceases receiving and retransmitting the signal, which decreases the overall signal coverage and signal strength. Because the signal has to be accessible from a helicopter, which can be flying anywhere within the coverage area, failure of one or more of the antennae constitutes a decrease in the performance of the entire system. The base station system is modelled as a k-out-of-n system, referred to as System 1. Fig. 1 provides an example of a k-out-of-n system, checked at scheduled non-periodic inspections (denoted by x_i) and opportunistically whenever n-k+1 components fail. The numbers 1,…,n-k above the black circles are denoting the ordinal number of component failures in the system over the time between inspections (and, hence, between failure rectifications). This is used to demonstrate an example of possible occurrence and accumulation of failures within a certain period of time.

Similarly to the above-stated assumptions for System 1, the following assumptions are made for System 2:

S.2.1: The number of soft-type components in System 2 is denoted as n_1, and the number of hard-type components is represented as n_2.

S.2.2: Similar to S.1.3, but with j=1,2,…,n_1+n_2 for System 2, where s=1,2,…,n_1 is the number of soft-type components and h=1,2,…,n_2 is the number of hard-type components in the system.

S.2.3: Opportunistic inspections are incurred whenever a hard failure occurs.

S.2.4: Hard failures create more opportunities for inspecting soft-type components and, consequently, influence the replacement and minimal repair numbers, as well as the downtime of components with soft failures (see Fig. 2).

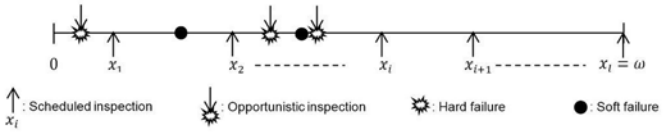


Fig.2: Sample hard and hidden soft failures, scheduled and opportunistic inspections within one life cycle for System 2.

Another problem involves optimising non-periodic inspection and maintenance policies for a MacPherson-type strut assembly found in automotive vehicles. Here, a shock absorber, a coil spring and a strut-to-mount nut are modelled as hard-type components, and lower and upper spring insulators, dust shield, jounce bumper and spring seat pad are modelled as soft-type components. The entire assembly constitutes a system and is referred to as System 2. Similarly to Fig. 1, Fig. 2 shows an example of hard and soft failures along with the scheduled and opportunistic inspections for System 2.

Once the end of the planning horizon (i.e. time ω) is reached, a new non-periodic optimal policy for maintenance and inspection can be established by repeating the outlined procedure. In the case of System 2, current ages of hard-type components can be taken into account when planning for the system life cycle.

The total cost of system maintenance and inspection is a metric used almost universally in different areas of industry for a large variety of systems. It is a convenient measure of the optimality of a system, since the latter's reliability and availability are connected through a range of costs, such as inspection and maintenance costs, component and system downtime penalties. For this reason, for both System 1 and System 2, the objective function is formulated based on the total cost of joint maintenance and inspection policy.

A closed deterministic formulation requires knowledge of all system parameters with certainty. However, this condition is not satisfied, because the failure ages of hidden soft-type components are unavailable. Instead, an expression is formulated to recursively find the expected system parameters (please refer to Appendix for details).

However, the recursive formula cannot be solved analytically because of some terms having multidimensional integrals requiring discretisation, which makes the computations cumbersome. Because of this, the present analysis is based on the results obtained from Monte Carlo simulations, as well as on the use of the genetic algorithm (GA).

To summarise, our objective is to find the non-periodic optimal maintenance policy m^* for System 1 and m_s^* for System 2, and optimal inspection policy x^*. The optimal joint policies are achieved through minimising for the whole system the total expected cost within the system's life cycle ω.

3. Model 1: joint optimisation of non-periodic inspection and corrective maintenance of *k*-out-of-*n* system with opportunistic inspections

In this section, we propose a model for a k-out-of-n system which may be non-periodically inspected at potential times $i\tau$, where τ is the minimal time unit. At the same time, maintenance optimisation is done for the discrete-valued number of component's minimal repairs until replacement. Overall, joint quasi-continuous and discrete optimisation is performed to obtain the joint optimal inspection and maintenance policies.

Maintenance optimisation is concerned with finding the best maintenance action in a particular system setting. In this paper, at each inspection point, the decision has to be made whether to minimally repair, or replace the failed component(s). Inspection optimisation then provides the best points in time at which these maintenance actions have to be taken in order to incur the lowest cost. Since failures are stochastic in nature, the total expected cost is used for optimality computations.

The scheduled non-periodic inspection policy $\boldsymbol{x}^a = (x_1, x_2, \ldots, x_l)$, where a refers to the inspection policy index, can be encoded as a binary sequence of '1's and '0's, where each '1' corresponds to a scheduled inspection and '0' corresponds to the lack thereof. Taking the number of digits of $\boldsymbol{x}^a$ to be l, each digit then corresponds to time $i\tau$. This binary representation lends itself naturally to the encoded "genome" strings used in the genetic algorithms, which makes it particularly convenient and effective for the purposes of inspection optimisation using the latter. The total number of possible distinct sched-

uled inspection policies is then 2^{l-1}, since there is always 1 inspection scheduled to occur at time ω. Hence, enumerating $\boldsymbol{x}^a$, $a = 1,2,\ldots,2^{l-1}$.

The components' failures are hidden, making their ages at failure unknown. For this reason, maintenance decisions cannot be age-based. Instead, the number of minimal repairs is counted for each component, and the decision of whether to minimally repair, or replace a failed component is based on the number of minimal repairs until replacement. Since all components are assumed to be identical in *k*-out-of-*n* configuration, only one optimal number of minimal repairs before replacement has to be found for a given system.

Based on observations S.1/2.11-S.1/2.13, it can be concluded that the component failure process is a counting process. Moreover, observations S.1/2.14-S.1/2.15 pertain to a Poisson process, and assumptions S.1.3 and S.2.2 further specify the sequence of random variables $\{\Phi(t)\}_{t\geq 0}$ describing the failure process as NHPP. Making use of assumptions S.1/2.9 and S.1/2.10,the expected number of failures $E[\Phi(\omega)]$ to time $\omega \geq 0$ is obtained as following:

$$E[\Phi(\omega)] = \Lambda(\omega) = \int_0^{\omega} \lambda(t)dt = \int_0^{\omega} \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} dt = \left(\frac{\omega}{\eta}\right)^{\beta}, \quad (1.1)$$

where $\Lambda(\omega)$ is a cumulative hazard function.

The actual number of failures is expected to fall within a $100*\alpha$ % confidence interval, with the upper confidence limit *UCL* given from Poisson distribution as:

$$UCL = \min\left\{\varphi \in \mathbb{N}_0 : P(\Phi(\omega) \leq \varphi) \geq 1 - \frac{1-\alpha}{2}\right\}$$
$$= \min\left\{\varphi \in \mathbb{N}_0 : \sum_{f=0}^{\varphi} P(\Phi(\omega) = f) \geq 1 - \frac{1-\alpha}{2}\right\}, \quad (2.1)$$

where *UCL* is the upper confidence limit for a component and is dependent on ω, and the probability of observing φ failures over planning horizon ω is given by $P(\Phi(\omega) = \varphi) = \frac{\left(E[\Phi(\omega)]\right)^{\varphi}}{\varphi!}\exp\left(-E[\Phi(\omega)]\right)$ for each soft-type component.

The total expected cost $E\left[C^{T,k,n}_{\boldsymbol{x}^a,m}\right]$ is formulated as following:

$$E\left[C^{T,k,n}_{\boldsymbol{x}^a,m}\right] = lc^I + F(\omega,\boldsymbol{t},f_s,m_c)c^{SD}$$
$$+\sum_{j=1}^{n}\left(c^M M_{\boldsymbol{x}^a,m,j}(\omega,t_j,f_s,m_c)\right.$$
$$+c^R R_{\boldsymbol{x}^a,m,j}(\omega,t_j,f_s,m_c)$$
$$\left.+c^D\left(\omega - U_{\boldsymbol{x}^a,m,j}(\omega,t_j,f_s,m_c)\right)\right), \quad (3)$$

where *l* is a number of scheduled inspections, c^I is a cost of a scheduled inspection, $\boldsymbol{t} = (t_1,t_2,\ldots,t_n)$ is a vector with initial ages of components, $F(\omega,\boldsymbol{t},k,n)$ is the expected number of failures for the system, c^{SD} is a downtime penalty for the system, c^D is a downtime cost per component per unit time, c^R is a per-component cost of corrective replacement, c^M is a per-component cost of minimal repair, t_j is an initial age of component j, m is a number of minimal repairs until replacement, $U_{\boldsymbol{x}^a,m,j}(\omega,t_j,f_s,m_c)$, $R_{\boldsymbol{x}^a,m,j}(\omega,t_j,f_s,m_c)$ and $M_{\boldsymbol{x}^a,m,j}(\omega,t_j,f_s,m_c)$ are, respectively, the expected uptime and the expected numbers of replacements and minimal repairs for each component. In the proposed formulation, expected values are generally obtained recursively using the generator function $G_{x_i,m_s,s}(\omega,t_s,f_s,m_c)$ with the variables indicated inside the brackets as parameters (see Appendix for details).

The results of the periodic optimisation from both exhaustive search and genetic algorithm search procedures were cross-verified and were found to be identical. Using the same logic and modifying the code to accommodate non-periodic frequency of inspections, we extrapolate the results to the non-periodic domain.

The optimal joint maintenance and inspection policies are determined by the optimal inspection policy $\boldsymbol{x}^*$ and the optimal number of minimal repairs until replacement m^*, respectively. Using calculations for the combinations of possible inspection and maintenance policies ($\boldsymbol{x}^a, m$), the optimal joint inspection policy ($\boldsymbol{x}^*, m^*$) can be obtained from searching for the smallest total expected cost as following:

$$(\boldsymbol{x}^*, m^*) = \min_{\boldsymbol{x}^a,m}\left\{E\left[C^{T,k,n}_{\boldsymbol{x}^a,m}\right]\right\},$$
$$\text{s. t.}: 0 \leq m \leq UCL,$$
$$x_i = \begin{cases} 1, \textit{if inspection occurs at time } i\tau \\ 0, \textit{if no inspection occurs at time } i\tau \end{cases},$$
$$i = 1,2,\ldots,l. \quad (4)$$

The expected values required for the calculation of $E\left[C^{T,k,n}_{\boldsymbol{x}^a,m}\right]$ are, however, unavailable for systems where some or all of components fail in hidden mode, because failure ages of these components are unavailable and cannot be formulated explicitly. This obstacle is overcome by using the simulation procedures described in Section 5.

4. Model 2: non-periodic inspection and corrective maintenance of hard-type and soft-type components with opportunistic inspection of soft-type components

This section describes the methodology for finding the optimal maintenance actions after failures and the optimal inspection policy for System 2, taking into account the fact that soft failures are hidden and the soft-type components' ages at the time of failure are unknown. The model resulting from this methodology is called "Model 2" further in the text.

Due to the different failure characteristics, hard-type components are analysed separately from the soft-type components. Hard failure times are known, since the system stops operating immediately when-

ever a hard failure occurs. The goal is to determine the optimal ages at which the hard-type components should be replaced, providing the lowest cost of inspection and maintenance for the entire system. In order to achieve this, the domain of possible replacement ages from which to choose the optimal ones has to be defined for each hard-type component. The replacement ages are represented by vector $\boldsymbol{\zeta}=\left(\zeta_1,\zeta_2,\ldots,\zeta_{n_2}\right)$, consisting of replacement ages for each hard-type component h=1,2,…,n_2 .

From the system life cycle's perspective, it is impractical to make the hard-type component's replacement age longer than the life span of the entire system, as represented by its planning horizon. It is assumed that an overhaul or similar renewal event is to take place at the end of the system's life cycle, at which point those hard-type components which have not been maintained over the system's operation will be replaced. Thus, the replacement ages for hard-type components can be assumed to be bounded by 0 from the bottom and a multiple of the system's life cycle length at the top, for example: $0<\zeta_h\leq 1.5\omega$, h=1,2,…,n_2 .The choice of the multiple of ω is arbitrary and depends on the practical considerations rather than the theoretical ones. The motivation for choosing 1.5 as a multiple for the upper bounds is to allow optimal replacement ages vary in the range greater than the system's life cycle length for greater generality, but at the same time not to waste computational resources checking for unrealistically long replacement ages that are impractical for planning purposes.

The optimal replacement ages for all hard-type components are represented by vector $\boldsymbol{\zeta}^*=(\zeta_1^*,\zeta_2^*,\ldots,\zeta_n^*)$ consisting of the optimal replacement ages for each hard-type component h=1,2,…,n_2.

Unlike those for hard failures, the soft failure times are unknown, which makes it impossible to base the optimisation procedure on the ages of soft-type components. Instead, maintenance decision can be based on the number of minimal repairs until replacement. Similarly to System 1 and using the same assumptions, the expected number of failures $E[\Phi_s(\omega)]$ for System 2 was obtained as following:

$$E\left[\Phi_s(\omega)\right]=\int_0^{\omega}\frac{\beta_s}{\eta_s}\left(\frac{t_s}{\eta_s}\right)^{\beta_s-1}dt=\left(\frac{\omega}{\eta_s}\right)^{\beta_s}. \tag{1.2}$$

The actual number of component failures, however, may vary, owing to the stochastic nature of component failures. Hard failures are assumed to be rectified immediately upon failure. Soft failures are rectified at the earlier of either a scheduled inspection, or hard failure (i.e. at opportunistic inspection). We may then get the general estimate on the upper bound of the number of minimal repairs until replacement from using Poisson distribution for $E[\Phi_s(\omega)]$ to construct a confidence interval at α level as following:

$$UCL_s=\min\left\{\varphi_s\in\mathbb{N}_0:\sum_{f_s=0}^{\varphi_s}P\left(\Phi_s(\omega)=f_s\right)\geq 1-\frac{1-\alpha}{2}\right\}, \tag{2.2}$$

where UCL_s is the upper confidence limit for soft-type component s and the rest of terms are as previously defined.

We assume that the number of minimal repairs before replacement m_S for soft-type component s, s=1,…,n_1 does not exceed the upper confidence limit UCL_s on the mean of component failures and may take on any value between 0 and UCL_s, inclusively. Thus, different cases are covered, ranging from replacing component at every failure to replacing it at (UCL_s+1)st failure, with m_C keeping track of the current number of minimal repairs. Furthermore, m_S thus selected serves as the criterion for making a maintenance decision. Component s is minimally repaired at each inspection so long as no more than m_S failures occur. It is then replaced on (m_S+1)st failure. The optimal number of minimal repairs until replacement m_s^* results in the lowest total expected cost $E\left[C^{T,HS}_{\boldsymbol{x}^a,m_s,\zeta_h}\right]$ for the entire system.

It should be noted, that unlike in preventive replacement models for mixed systems composed of hard- and soft-type components encountered in Babishin and Taghipour [4], corrective replacement models may exclude the costs of hard-type components from the optimisation. This is because the hard-type components are replaced at the optimal replacement ages if they fail, and not at the scheduled inspection times when they are still operational. This makes the optimal non-periodic inspection independent of the costs of hard failures.

In order to obtain the lowest expected cost, all combinations of non-periodic inspection schedules, the numbers of minimal repairs before replacement, and various ages as threshold for replacement have to be considered for all components. The expected costs thus calculated can then be searched for the lowest value. However, the size of the search space is very large in this case. For this reason, this problem, albeit in the context of periodic inspections, has been previously broken down into several stages for maintenance and inspection optimisation for all hard-type components in Stage 1, marginal optimisation of the maintenance decision for each soft-type component in Stage 2 and optimisation of the inspection period for the entire system in Stage 3, using Monte Carlo simulation for marginal multi-stage optimisation [4].

In the present paper, global system-level optimisation is performed, which requires simultaneous optimisation of all decision variables. This results in a dramatic increase of the search space. The latter is greatly reduced by means of the genetic algorithm. This allows optimising for both inspection and maintenance jointly in one stage. The total expected cost $E\left[C^{T,HS}_{\boldsymbol{x}^a,m_s,\zeta_h}\right]$ is calculated as following:

$$\begin{aligned}E\left[C^{T,HS}_{\boldsymbol{x}^a,m_s,\zeta_h}\right]=lc^I+\sum_{s=1}^{n_1}\Big(&c_s^M M_{\boldsymbol{x}^a,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)\\&+c_s^R R_{\boldsymbol{x}^a,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)\\&+c_s^D\left(\omega-U_{\boldsymbol{x}^a,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)\right)\Big)\end{aligned} \tag{5}$$

where superscript HS indicates the cost for System 2 consisting of components with both soft and hard failures, $\theta=(\theta_1,\theta_2,\ldots,\theta_h,\ldots,\theta_{n2})$ is a vector containing the initial ages of hard-type component h, h=1,2,…,n_2, $\boldsymbol{\zeta}=\left(\zeta_1,\zeta_2,\ldots,\zeta_{n_2}\right)$ is a vector with replacement ages of hard-type components, t_S is the initial age of component s, c_s^M is a minimal repair cost of component s, c_s^R is a replacement cost for each soft-type component, c_s^D is a cost of downtime for component s, m_C is the current number of minimal repairs, and terms $U_{\boldsymbol{x}^a,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)$, $R_{\boldsymbol{x}^a,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)$ and $M_{\boldsymbol{x}^a,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)$ represent, respectively, the expected uptime and the expected numbers of replacements and minimal repairs for each soft-type component s, s=1,2,…,n_1.

Using the calculations for the combinations of possible inspection and maintenance policies $\left(\boldsymbol{x}^a,m_s,\zeta_h\right)$, the optimal joint inspection and maintenance policy $\left(\boldsymbol{x}^*,m_s^*,\boldsymbol{\zeta}^*\right)$ can be obtained from searching for the smallest total expected cost as following:

$$\left(\boldsymbol{x}^*, m_s^*, \zeta^*\right) = \min_{\boldsymbol{x}^a, m_s, \zeta_h} \left\{ E\left[C^{T,HS}_{\boldsymbol{x}^a, m_s, \zeta_h} \right] \right\},$$
$$\text{s. t.}: 0 \le m_s \le UCL_s,$$
$$s = 1,2,\ldots,n_1,$$
$$0 \le \theta_h \le \zeta_h,$$
$$h = 1,2,\ldots,n_2,$$
$$x_i^a = \begin{cases} 1, \text{ if an inspection occurs at time } i\tau \\ 0, \text{ if no inspection occurs at time } i\tau \end{cases}$$
$$i = 1,2,\ldots,l,$$
$$a = 1,2,\ldots,2^{l-1}. \quad (6)$$

The following section outlines the general simulation procedure used for optimisation.

5. Simulation Model

Simulation procedure is similar for both systems, but differs in some details as a result of the difference in the types of system's components.

5.1. Simulation Model for *k*-out-of-*n* System (System 1)

The simulation for the *k*-out-of-*n* system takes as inputs the values of x_i^a, m, k, n, ω, τ, t_j, β, η, c^M, c^{CR}, c^{PR}, c^D, c^I and c^{SD}. Let the random variable Y_j (uptime of component j) have a Weibull distribution with parameters β and η. This component has an age t_j and probability that the time-to-failure is equal to χ_j, which is given by formula:

$$P\left(\mathrm{Y}_j = \chi_j + t_j | \mathrm{Y}_j \ge t_j\right) = \frac{\mathcal{R}\left(\chi_j + t_j, \beta, \eta\right)}{\mathcal{R}\left(t_j, \beta, \eta\right)}, \quad (7)$$

where $\mathcal{R}\left(t_j, \beta, \eta\right)$ is a reliability function. To generate the time-to-failure for component j, $j = 1,2,\ldots,n$, we first generate a random number b, which has a uniform distribution on interval $[0;1]$, and next calculate a quantile of order b for conditional distribution in Equation (7).The time-to-failure χ_j is generated as following:

$$\chi_j = \eta\left(\left(\frac{t_j}{\eta}\right)^{\beta} - \ln(b)\right)^{1/\beta} - t_j. \quad (8)$$

Generated times-to-failure are then compared with the time $i\tau$ of the earliest scheduled inspection flagged as '1' in $\boldsymbol{x}$. While $\chi_j < i\tau$, the number of failures of component j is increased by 1. Once total component failures in the system accumulate to $n-k+1$ failures, failure of entire system occurs, giving rise to opportunistic inspection, during which all failures are discovered. A failed component is minimally repaired if it has failed for m times or fewer; alternatively, it is replaced, and the failure count for it is reduced to zero.

While $\chi_j < i\tau$, but there are fewer than $n-k+1$ failed components, the latter are fixed at the following scheduled inspection. Again, a failed component is minimally repaired if it has failed for m times or fewer; alternatively, it is replaced, and the failure count for it is reduced to zero.

While $\chi_j > i\tau$, the simulation clock is moved forward to the inspection time, since there is no failed component to be discovered at inspection.

The simulation clock is updated at the times of events, such as component failures, system failures, and scheduled and opportunistic inspections. The downtime of component j, the number of system failures, the number of minimal repairs and replacements are all updated at each event's time as well. At the same time, the ages of the surviving components, the time until the next failure, and the time until the scheduled inspection are also updated. The simulation stops when the system's life cycle is completed. Running the simulation for a large number of times provides the expected values of the random variables F, $M_{\boldsymbol{x}^a,m,j}$, $R_{\boldsymbol{x}^a,m,j}$ and $U_{\boldsymbol{x}^a,m,j}$.

A given policy ($\boldsymbol{x}^a, m$) prescribes the choice of the maintenance action at each simulation run. Varying the values of $\left(\boldsymbol{x}^a, m\right)$ in the ranges $\boldsymbol{x}^a = (0,0,\ldots,0),(1,0,\ldots,0),\ldots,(1,1,\ldots,1)$, $0 \le m \le UCL$, the total expected cost $E\left[C^{T,k/n}_{\boldsymbol{x}.m}\right]$ for policy $\left(\boldsymbol{x}^a, m\right)$ is computed and saved. Thus, the total number of distinct policies in the decision space for Model 1 is $2^{\omega/\tau - 1} * UCL$.

Lastly, the joint optimal inspection and maintenance policy $\left(\boldsymbol{x}^*, m^*\right)$ is found from searching for the minimum $E\left[C^{T,k,n}_{\boldsymbol{x}^*.m^*}\right]$.

5.2. Simulation model for a system with hard-type and soft-type components (System 2)

The general simulation procedure for System 2 is similar to that described for System 1. The following input variables are used: x_i^a, m_c, m_s, ω, τ, t_s, $\theta_h, \beta_s, \beta_h, \eta_s, \eta_h, \zeta_h$, c_s^M, c_h^D, c_s^R, c_h^D, c_s^D and c^I. The same procedure as discussed in the previous section is used to generate the times for events and update the simulation clock.

If $\chi_s < \chi_h < i\tau$ for the generated soft failure time χ_s, the closest hard failure time χ_h and the closest non-periodic inspection time $i\tau$, then an opportunistic inspection is occurring at the closest hard-type component's failure time. The soft-type component's failure is detected at this moment, and if the total number of previous failures is less than m_s, the component is minimally repaired; otherwise, it is replaced, and its failure count is reset to zero.

When $\chi_h \le \chi_s < i\tau$, soft failure is fixed at the soonest scheduled inspection.

If $\chi_h < i\tau$, the age of the failed hard-type component is compared with its corresponding replacement age ζ_h, and the hard-type component is replaced if $\theta_h > \zeta_h$, or it is minimally repaired otherwise. The component's age is set to zero at replacement.

Changing the values of a given joint inspection and maintenance policy $(\boldsymbol{x}^a, m_s, \zeta_h)$ in the ranges $\boldsymbol{x}^a = (0,0,\ldots,0),(1,0,\ldots,0),\ldots,(1,1,\ldots,1)$, $0 \le m_s \le UCL_s$, $s = 1,2,\ldots,n_1$, $0 \le \zeta_h \le 1.5\omega$, $h = 1,2,\ldots,n_2$, the total expected cost for each policy is computed and saved. Thus, the total number of distinct policies in the decision space for Model 2 is $2^{\omega/\tau-1} * UCL_{n_1}{}^{n_1} * (1.5\omega/\tau)^{n_2}$.

Finally, the optimal solution is the one with the minimum cost $E\left[C^{T,HS}_{\boldsymbol{x}^*,m_s^*,\zeta_h^*}\right]$ over all input variables, i.e. inspection policies, possible values of the number of minimal repairs before replacement for soft-type components and possible values for the optimal replacement ages for hard-type components. The resulting triple $\left(\boldsymbol{x}^*, m_s^*, \zeta^*\right)$ represents the optimal joint inspection and maintenance policy.

In using simulation as described above, however, there is a significant drawback related to the search method used for optimisation. The search method for the optimal joint policy is highly sensitive to the size of the problem's search space, which, in turn, is related to the choice of τ, UCL_s, etc. Therefore, as the number of components and/or the life cycle length are increased, a dramatic increase is also observed in both the search space and the simulation time. The complexity of the problem also increases with a decrease of τ, as in this case the number of non-periodic inspections increases, and the possible number of inspection policies quickly explodes. Thus, based on all of these, a reduction of computational complexity and an increase in computational efficiency are required.

Reducing the search space and problem complexity can be achieved by decreasing the number of instances of calculating the total expected cost. The genetic algorithm provides a powerful heuristic search means to do this. The proposed approach is further discussed below.

6. Numerical Examples

The present section provides examples for each of the models developed in the preceding sections.

6.1. Model 1: k-out-of-n system with opportunistic inspections

We first consider a 3-out-of-5 redundant system with all components in the "as-good-as-new" state and parameters given in Table 1.

Table 1. Parameters and Costs of the Power Law Intensity Function.

Case	β	η (months)	Minimal repair cost, c^M	Replacement cost, c^R	Component downtime cost, c^D	System downtime cost, c^{SD}	Fixed inspection cost, c^i
1	1.5	3.5	$75	$200	$60	$550	$50
2	1.5	3.5	$75	$200	$80	$550	$50
3	1.5	3.5	$75	$200	$60	$350	$50
4	1.5	3.5	$75	$200	$60	$550	$100
5	1.5	5	$75	$200	$60	$550	$50

All models were implemented in MATLAB, and "ga" function was called for the genetic algorithm calculations. We took 200 as the number of individuals in a population of candidate solutions and 1 as the number of individuals guaranteed to survive to the next generation ("elite count" parameter). We used gene-based coding scheme, where each element in the coded "genome" string could change independently. Each coded string represented an inspection policy, where the string's length corresponded to the number of possible inspections within the planning horizon less one (since an inspection is always done at the end of the planning horizon according to assumption S.1/2.3), which was 11 for the planning horizon of 1 year with a monthly inspection frequency. Binary alphabet was used for coding, with '1' denoting inspection and '0' denoting no inspection, as mentioned previously. Crossover was performed by taking the weighted average of the parents *parent1* and *parent2* according to the formula:

$$child = parent1 + rand * (parent2 - parent1), \quad (9)$$

where *rand* is a random number. The children thus produced are then within the hypercube defined by the parents at the opposite vertices. The crossover fraction, which is the fraction of the next generation (apart from the elite individuals) produced by crossover, was taken at the default value of 0.8. Adaptive feasible mutation was used, which randomly generated directions that were adaptive toward the last generation. A step length of 1 was chosen along each direction, so that linear constraints and bounds were satisfied. Taking into consideration the running time of the simulation and the error in the calculated mean values, the simulation procedure was repeated for 5000 runs for each of the possible inspection policies with the planning horizon of 12 months and τ = 1 month. The values of the calculated total expected costs were saved and then averaged for each inspection policy. Optimisation was then performed using genetic algorithm, and the output of the GA for Case 1 is given in Fig. 3.

For the GA search procedure, the highest fitness value is equivalent to the lowest total expected cost. The genetic algorithm search used a limit of 100 generations, a stall generation limit of 50, an elite count of 1 and the tolerance limit of 10^{-5}. We repeated the GA for 30 trials with different seeds and chose the best solution with the lowest total expected cost.

As can be seen from Fig. 3, the best solution was found after 11 generations. The top of Fig. 3 plots the best and the mean values of the penalty function for each generation. Because the formulation is integer problem with constraints, the penalty function includes a term for infeasibility. If the generation results in a feasible solution, the penalty function is identical to the fitness function, which is the total expected cost for each generation. Otherwise, the penalty function is the maximum fitness function among the feasible generations plus the sum of the constraint violations of the infeasible points. This ensures that an infeasible solution is not selected in the optimisation process.

The fitness function approaches optimality in the sense of the total expected cost with the generations' number increasing, and converges after about 40 generations. The global optimality has been verified for the periodic optimisation of the *k*-out-of-*n* system by comparing the results from both exhaustive search and genetic algorithm search procedure, which have been found to be identical. Moreover, we found the costs to be very close, so that any locally optimal cost is not far from the globally optimal result obtained from the simulation.

The middle of Fig. 3 shows the average nearest neighbour distances for each generation. Generally, lower distance implies a more localised search after 40 generations.

The bottom of Fig. 3 displays the best, worst and mean fitness function scores. One can notice that starting at generation 39, the dis-

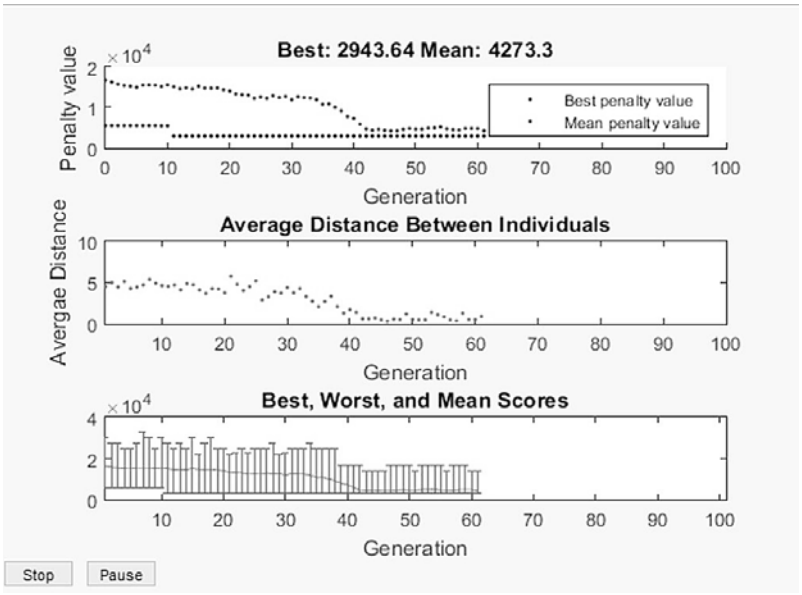


Fig. 3. Genetic algorithm's output and results for Case 1 of a 3-out-of-5 system

tance between the best and the worst scores (i.e. total expected costs) decreases, while the mean score approaches the best score.

Baseline Case 1 was compared with the other cases for sensitivity analysis. The results of the genetic algorithm's search procedure for the cases are given in Table 2. Note that the last column in the table represents the number of distinct inspection policies out of 2048 inspection policies. The fact that optimal policies were found well before reaching the total number of policies suggests fast convergence of the genetic algorithm in the neighbourhood of the optimal cost. Such quick convergence toward the optimal solution is likely resulting from the smoothness of search space in vicinity of the optimal solution.

Table 2 shows that the optimal inspection policy suggests that for the baseline Case 1 system, 9 inspections should be performed at month 1, 3, 4, 6, 7, 8, 9, 11 and 12 within the 12-months planning horizon, and the optimal maintenance policy suggests that failed components should be replaced upon failure after having been minimally repaired for 5 times.

For the system with the component downtime cost of \$80 (Case 2), 9 inspections should be done at month 2, 3, 5, 6, 7, 9, 10, 11 and 12, and the optimal maintenance policy prescribes component replacement on 7th failure within the 12-months planning horizon. Unsurprisingly, the total expected cost for Case 2 is greater than that for Case 1, since the cost of component downtime is greater for Case 2 than for Case 1.

For Case 3, the lower downtime cost for the system results in lower $E\left[C^{T,k,n}_{\boldsymbol{x}^*,m^*}\right]$, fewer inspections and greater m^* compared with Case 1. Lower system downtime penalty translates into more of allowable system downtime, which necessitates fewer inspections (7 for Case 3 vs. 9 for Case 1) and a much greater m. The fact that $m^* = UCL$ implies that it is economically infeasible to replace failed components when the system downtime penalty is significantly decreased. However, the total effect of the decrease of the system downtime penalty by \$200 is reduced by the increased downtime as a result of the fewer inspections and greater component deterioration due to fewer replacements, all of which are reflected in the total expected cost's decrease of only \$168.01.

The optimal joint inspection and maintenance policies for Case 4 are close to those of Case 3, but the total expected cost is greater than that for either Case 3 or Case 1. Unsurprisingly, increasing the cost of system inspection results in the increase in $E\left[C^{T,k,n}_{\boldsymbol{x}^*,m^*}\right]$. Removing the effect of the total cost of inspection, it can be seen that the remaining expected cost for Case 4 is lower than that for Case 1 by \$70.47. This is likely the result of the 2-fold increase in the optimal number of minimal repairs before replacement, which results in fewer component replacements prescribed by Case 4 compared with those for Case 1.

Finally, for Case 5, the total expected cost is the lowest among all the tested cases. This can be explained by higher scale (spread) parameter of the time-to-failure distribution, which implies fewer failures within the same time interval for Case 5 compared to the other cases. Using Equation (11) from Babishin and Taghipour [3], the calculated expected number of system failures is approximately 5.7 for Cases 1-4 and only about 2.6 for Case 5 – a decrease by over 121 % for Case 5 compared to the other cases. This also results in the fewest optimal number of inspections (6) among all the cases and, also, a slightly higher optimal number of minimal repairs until replacement (7) and, correspondingly, fewer component replacements compared to that for Case 1.

6.2. Model 2: system with hard-type and soft-type components and opportunistic inspections

We consider a system composed of m_1=5 components prone to hidden soft failure and m_2=3 components prone to hard failure, all of which are initially "as-good-as-new". Two cases are considered: Case 1 (Baseline) and Case 2 (1.5-time greater monthly downtime penalty cost compared to Baseline).The input parameters for the failure distributions, the costs of minimal repair, replacement and downtime are given for both cases in Table 3.

The fixed cost of scheduled inspections c_I=\$25. Both cases are simulated for 1000 runs with the planning horizon of 12 months and τ = 1 month. Lower number of simulation runs had to be used because of the much greater time required to run the simulation and the genetic algorithm for the hard-and-soft-type system compared to a *k*-out-of-*n* system, which is a result of the former's greater complexity.

The output of the GA for Case 1 is given in Fig. 4. The genetic algorithm search used a limit of 100 generations, a stall generation limit of 50, an elite count of 1 and the tolerance limit of 10^{-5}.

As can be seen from the top graph in Fig. 4, the best solution ("best penalty value") was found after 71 generations.The global optimality has been verified by comparing the results from both exhaustive search and genetic algorithm search procedure, which have been found to be identical.

Table 2. Optimal policies from the genetic algorithm

Case #	Distinction from baseline	Total expected cost, $E\left[C^{T,k,n}_{\boldsymbol{x}^*,m^*}\right]$	Optimal inspection policy, $\boldsymbol{x}^*$	Optimal maintenance policy, m^*	Number of inspection policies analysed, a
1	—	\$2943.64	(1 0 1 1 0 1 1 1 1 0 1 1)	5	1755
2	c^D = \$80	\$3098.46	(0 1 1 0 1 1 1 0 1 1 1 1)	6	1654
3	c^{SD} = \$350	\$2775.63	(1 0 1 0 1 1 0 1 1 0 0 1)	11	1709
4	c^i = \$100	\$3123.17	(1 0 1 1 0 1 0 1 0 1 0 1)	10	1638
5	η = 5	\$1946.68	(0 1 0 0 1 1 0 0 1 1 0 1)	7	1104

Table 3. Power law intensity parameters and costs for different components of Case 1 (Baseline) and Case 2.

Component type		β_j	η_j (months)	Minimal repair cost, c_j^M	Replacement cost, c_j^R	Case 1 downtime penalty cost/ month, c_j^D	Case 2 downtime penalty cost/ month, c_j^D
Soft	1	1.3	3.5	\$70	\$200	\$80	\$120
	2	2.8	4.6	\$45	\$150	\$55	\$82.5
	3	2.1	2.7	\$100	\$300	\$85	\$127.5
	4	3.2	7.0	\$75	\$240	\$90	\$135
	5	1.7	3.6	\$125	\$325	\$100	\$150
Hard	1	1.5	8.7	–	–	–	–
	2	1.2	6.0	–	–	–	–
	3	1.7	7.5	–	–	–	–

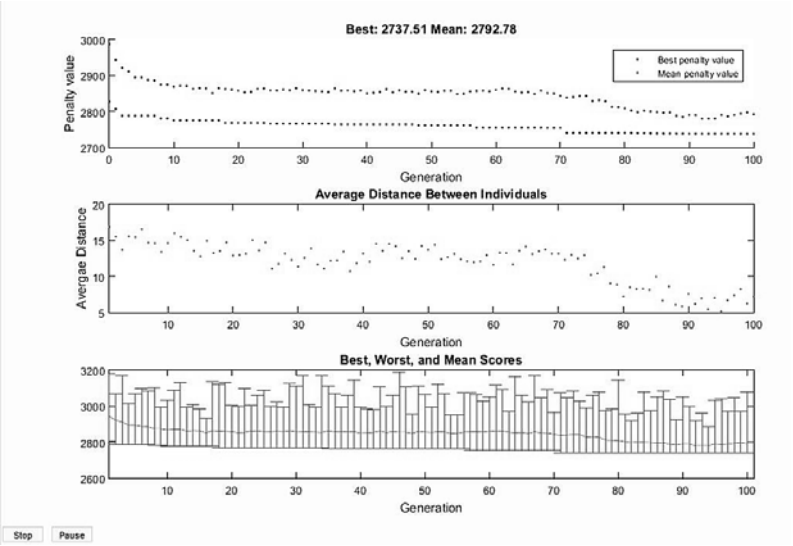


Fig. 4. Genetic algorithm's output and results for Case 1 of a system composed of soft- and hard-type components.

The graph in the middle of Fig. 4 shows the nearest neighbour distances for each population member. Overall, lower distance implies the search space narrowing down starting at around the 80th generation on.

The graph at the bottom in Fig. 4 contains the best, worst and mean fitness function scores. One can see the mean scores of the fitness function are minimal at generations 92-94, owing to the lowest worst scores, while the best scores remain unchanged starting at generation 71.

The soft-type and the hard-type component parameters from Table 3 are used to optimise for the joint inspection and maintenance policies using genetic algorithm. The GA procedure was repeated for 5 trials, and the best results are shown. The results are obtained for both Case 1 and Case 2. The optimal maintenance policies, i.e. the optimal numbers of minimal repair until replacement and the optimal replacement ages, are provided for both cases in Table 4.

As can be seen from Table 4, the optimal replacement ages for some of the hard-type components (namely, for hard-type component 1 for Case 1 and hard-type components 1 and 3 for Case 2) exceed the planning horizon. This simply means that these components would be replaced only at the end of the planning horizon and would be minimally repaired if they fail at any time until then. Also, changes in the optimal maintenance policies for components with soft and hard failures suggest that they are affected by changes in the component downtime penalty.

The resultant optimal inspection policy was also found for both cases to be as shown in Table 5. As the last column of Table 5 suggests, 9935 out of 10240 inspection policies were checked in order to find the optimal one. The timing results suggest that, generally, in the case of non-periodic inspections, validating GA procedure by checking all possibilities is not practical. The advantage of using GA is further supported by Hajipour and Taghipour, who previously found that a genetic algorithm required only about 7 % of the time needed for exhaustive search [16].

Table 4. Optimal maintenance policies for Case 1 and Case 2

Case #	Distinction from baseline	Soft-type component, s	Optimal number of minimal repairs before replacement, m_s^*	Hard-type component, h	Optimal replacement age, ζ_h (months)
1	––	1	7	1	14.46
		2	1	2	8.25
		3	6	3	11.83
		4	8	–	–
		5	13	–	–
2	c^D(Case 2) = 1.5 cD (Case 1)	1	9	1	16.61
		2	1	2	6.77
		3	3	3	13.98
		4	0	–	–
		5	10	–	–

The optimal inspection policy for Case 1 implies that the system is inspected 4 times in months 5, 8, 10 and 12. The optimal inspection policy for Case 2 is drastically different with 10 inspections occurring on a monthly basis in months 3–12. Thus, as a result of a 1.5-time increase in the per-component monthly downtime penalty cost, the optimal inspection policy alone for Case 2 costs \$150 more (\$25 · (10−4) inspections) than that for Case 1. This leaves another \$960.90 as the increase in the cost of the optimal maintenance policy out of the total increase of \$1110.90 (\$3848.41-\$2737.51) in the total expected cost for Case 2 compared to Case 1. Converting dollars into percent-

Table 5. Optimal policies from the genetic algorithm

Case #	Total expected cost, $E\left[C^{T,HS}_{\mathbf{x}^a,m_s,\zeta_h}\right]$	Optimal inspection policy, $\mathbf{x}^*$	Number of inspection policies analysed, a
1	\$2737.51	(0 0 0 0 1 0 0 1 0 1 0 1)	9935
2	\$3848.41	(0 0 1 1 1 1 1 1 1 1 1 1)	9935

Table 6. Absolute and relative changes in the expected costs for Case 2 compared to Case 1.

Change (Case 2 – Case 1)	Average system downtime penalty cost	Average optimal replacement ages	Average optimal number of minimal repairs before replacement	Optimal inspection policy cost	Optimal maintenance policy cost	Total expected cost, $E\left[C^{T,HS}_{\boldsymbol{x}^a,m_s,\zeta_h}\right]$
Absolute	+ \$205	+ 0.94	- 2.40	+ \$150	+ \$960.90	+ \$1110.90
Relative	+ 50 %	+ 8.13 %	- 52.17 %	+ 5.48 %	+ 35.10 %	+ 40.58 %

ages, the total expected cost is over 40 % greater for Case 2 compared for Case 1. A large reduction in m_s^* for 3 of 5 soft-type components is meant to decrease these components' downtime by reducing their ages to 0 with each replacement more frequently for Case 2 than for Case 1. Averaged across all soft-type components, m_s^* =4.6 for Case 2 and m_s^* =7 for Case 1 – a decrease of over 52 % as a result of the downtime costs increasing by 50 %. At the same time, the optimal replacement ages of the hard-type components increased, on average, by only slightly over 8 % (which is less than τ) for Case 2 compared to Case 1. These results are summarised inTable 6.

As can be concluded from Table 6, the average optimal number of minimal repairs before replacement is most sensitive to change in the downtime penalty, followed by the total expected cost and the optimal maintenance policy cost, with an increase in downtime penalty causing a decrease in m_s^* and increase in each of $E\left[C^{T,HS}_{\boldsymbol{x}^a,m_s,\zeta_h}\right]$ and the optimal maintenance policy cost. On the contrary, the optimal inspection policy cost is the least sensitive, followed by the average optimal replacement ages, where an increase in downtime penalty increases each of the optimal inspection policy cost and the average optimal replacement ages.

7. Conclusions

In the present article, optimisation of non-periodic maintenance and inspection was considered for complex multicomponent systems with either *k*-out-of-*n* redundant configuration, or with components prone to hard and soft failures. Aside from scheduled inspections, components can be checked opportunistically at system failures. Making the unit of time sufficiently small allows to treat the planning horizon as quasi-continuous for possible non-periodic inspections, which gives a much greater flexibility and variety in the choice of available inspection policies at the expense of computational complexity, when compared to the periodic inspections. Since soft failures are hidden, component's age cannot be used as the criterion for maintenance optimisation. Instead, maintenance policies are defined by the number of minimal repairs before replacement for each component prone to hidden soft failure. The optimal policies are then found by jointly optimising the inspection and maintenance policies for the lowest total expected cost. Using simulation and a genetic algorithm to implement the joint optimisation was found to be an efficient and convenient method to find the optimal policies for large and complex systems. This appears to be a promising method for optimisation with regards to complex systems with multiple decision parameters.

Acknowledgements

The authors acknowledge the financial support from the Natural Sciences and Engineering Research Council (NSERC) of Canada for this research.

Appendix

This section presents the recursive functions for calculating the expected values in the simulation. The general principle of constructing a generator function is to calculate the expectation of an event of interest, such as a failure, minimal repair, replacement, or uptime. Since components are not independent of each other, and their survival is related to survival of other components, it is necessary to condition their probabilities on reliabilities and failures of other components. For example, replacement of a soft-type component is conditional on its (UCL_s+1)st failure and inspection, which itself is conditional on the failure of any hard-type component and inspection policy. In order to make the generator function general and applicable to different categories of components and events, placeholder functions are used extensively. Depending on the placeholder function $\psi(y,z,I_s)$, either $M_{\boldsymbol{x}^a,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)$, $R_{\boldsymbol{x}^a,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)$, or $U_{\boldsymbol{x}^a,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)$ functions are obtained. For $\psi(y,z,I_s)$, I_s is 0, 1, or 2, if component s is, respectively, minimally repaired, replaced, or has not failed in an interval. The first failure times of soft-type component s and the hard subsystem are, respectively, random variables Y and Z, and their density and reliability functions are $f^Y(y\,|\,t_s)$, $f^Z(y|\boldsymbol{\theta}),\mathcal{R}^Y(y|t_s)$ and $\mathcal{R}^Z(z|\boldsymbol{\theta})$. We then formulate the generator function for the expected uptime and the expected numbers of replacements and minimal repairs for each component over inspection interval $[0,i\tau]$ by conditioning on the random variables $Y=y$ and $Z=z$. Taking $q^{k,n}(z)$ to be the probability that the failure of the system at time $Z=z$ is due to the failure of component s:

$$q^{k,n}(z)f^Z(z|t_s)=\lambda_s(z\,|\,t_s)\exp\left(-\int_0^z\lambda^H(\chi\,|\,t_s)d\chi\right).$$

We start by assuming only one scheduled inspection at the end of the planning horizon and work backward to initial time 0. For the system with soft-type and hard-type components, the expected number of minimal repairs for a soft-type component can be calculated as following:

$$G_{x_1,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)=\int_0^{i\tau}\int_0^{z}\sum_{h=1}^{n_2}\Big\{\Big(\Big[\Big[\psi(y,z,0)+G_{x_1-z,m_s,s}\left(\omega-z,t_s+y,\boldsymbol{\theta}\oplus z,f_s+1,\boldsymbol{\zeta}^*,m_c+1\right)\Big]r_s^Y(m_c,m_s)$$
$$+\Big[\psi(y,z,1)+G_{x_1-z,m_s,s}\left(\omega-z,0,\boldsymbol{\theta}\oplus z,0,\boldsymbol{\zeta}^*,0\right)\Big]\bar{r}_s^Y(m_c,m_s)\Big)r^{k,\mathbf{n}}(f_s,k)$$
$$+\Big(\Big[\psi(y,z,0)+G_{x_1-z,m_s,s}\left(\omega-z,t_s+y,(\boldsymbol{\theta}\oplus z)^{(0_h)},f_s+1,\boldsymbol{\zeta}^*,m_c+1\right)\Big]r_s^Y(m_c,m_s)$$
$$+\Big[\psi(y,z,1)+G_{x_1-z,m_s,s}\left(\omega-z,0,(\boldsymbol{\theta}\oplus z)^{(0_h)},0,\boldsymbol{\zeta}^*,0\right)\Big]\bar{r}_s^Y(m_c,m_s)\Big)\bar{r}^{k,n}(f_s,k)\Big\}q^h(z)f^Y(y\,|\,t_s)f^Z(z|t_s)\,dydz$$
$$+\int_0^{i\tau}\sum_{s=1}^{n_1}\Big\{\Big[\psi(z,z,2)+G_{x_1-z,m_s,s}\left(\omega-z,t_s+z,\boldsymbol{\theta}\oplus z,f_s,\boldsymbol{\zeta}^*,m_c\right)\Big]r^{k,n}(f_s,k)$$
$$+\Big[\psi(z,z,2)+G_{x_1-z,m_s,s}\left(\omega-z,t_s+z,(\boldsymbol{\theta}\oplus z)^{(0_h)},f_s,\boldsymbol{\zeta}^*,m_c\right)\Big]\bar{r}^{k,\mathbf{n}}(f_s,k)\Big\}q^{k,n}(z)\mathcal{R}^Y(z\,|\,t_s)f^Z(z|t_s)\,dz$$
$$+\int_0^{i\tau}\Big[\psi(y,i\tau,0)r_s^Y(m_c,m_s)+\psi(y,i\tau,1)\bar{r}_j^Y(m_c,m_s)\Big]f^Y(y\,|\,t_j)dy\mathcal{R}^Z(i\tau|\boldsymbol{\theta})$$
$$+\psi(i\tau,i\tau,2)\mathcal{R}^Y(i\tau\,|\,t_j)\mathcal{R}^Z(i\tau|\boldsymbol{\theta}),\tag{A.1}$$

where:

$$r_s^Y(m_c,m_s)=\begin{cases}1, & if\ m_c\le m_s\\ 0, & otherwise\end{cases},$$

$$r^{k,n}(f_s,k)=\begin{cases}1, if\ f_s<n-k\\ 0, otherwise\end{cases},$$

$$\bar{r}_s^Y=1-r_s^Y,\ \bar{r}_h^Z=1-r_h^Z.$$

The placeholder function $\psi(y,z,I_s)$ varies depending on the random variable of interest.

For the expected number of minimal repairs:

$$\psi(y,z,I_s)=\begin{cases}1, if\ I_s=0\\ 0, otherwise\end{cases}.$$

For the expected number of replacements:

$$\psi(y,z,I_s)=\begin{cases}1, if\ I_s=1\\ 0, otherwise\end{cases}.$$

For the expected uptime:

$$\psi(y,z,I_s)=\begin{cases}y, if\ \ y<z\\ z, otherwise\end{cases}.$$

These placeholder functions modify the generator function $G_{x_1,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)$, which is used for the first inspection (initial state), and then, through recursion, in $G_{x_i,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\boldsymbol{\zeta},m_c)$ (see below and Equation (A.2)). Thus, the values of uptime and number of minimal repairs before replacement are obtained from Monte Carlo simulation. For a particular component in one run, its expected uptime is obtained from simulating this component in conjunction with the other components for the length of the entire planning horizon, recording its downtime, and then subtracting its downtime from the planning horizon to obtain its uptime. This can be repeated and averaged to obtain the component's expected uptime.

For the number of minimal repairs before replacement, the upper confidence limit on the expected number of component failures is obtained from Equations 1.1-2.1, or 1.2-2.2 for each soft-type component. Simulations are then run for all values of m or m_s, starting from 0 and up to and including UCL or UCL_s.

$\boldsymbol{\theta} \oplus z$ denotes addition of a scalar z to each of the coordinates of vector $\boldsymbol{\theta}$, i.e. $\boldsymbol{\theta} \oplus z = \left(\theta_1 + z, \theta_2 + z, \ldots, \theta_{n_2} + z\right)$, and $\left(\boldsymbol{\theta} \oplus z\right)^{(0_h)}$ means the h^{th} coordinate of vector $\boldsymbol{\theta} \oplus z$ is replaced by zero, i.e. $\left(\boldsymbol{\theta} \oplus z\right)^{(0_h)} = \left(\theta_1 + z, \ldots, \theta_{h-1} + z, 0, \theta_{h+1} + z, \ldots, \theta_{n_2} + z\right)$.

Equation (A.1) is extended to obtain the expected value of a random variable of interest for inspection policy x_i when inspections are performed at $i\tau$:

$$
\begin{aligned}
G_{x_i,m_s,s}(\omega,t_s,\boldsymbol{\theta},f_s,\zeta,m_c) = \int_0^{i\tau}\int_0^{z}\sum_{h=1}^{n_2}\Big\{\Big(\Big[\psi(y,z,0) + G_{x_i-z,m_s,s}\left(\omega-z,t_s+y,\boldsymbol{\theta}\oplus z,f_s+1,\zeta^*,m_c+1\right)\Big] r_s^Y(m_c,m_s) \\
+\Big[\psi(y,z,1) + G_{x_i-z,m_s,s}\left(\omega-z,0,\boldsymbol{\theta}\oplus z,0,\zeta^*,0\right)\Big]\bar{r}_s^Y(m_c,m_s)\Big) r^{k,\mathbf{n}}(f_s,k) \\
+\Big(\Big[\psi(y,z,0) + G_{x_i-z,m_s,s}\left(\omega-z,t_s+y,(\boldsymbol{\theta}\oplus z)^{(0_h)},f_s+1,\zeta^*,m_c+1\right)\Big] r_s^Y(m_c,m_s) \\
+\Big[\psi(y,z,1) + G_{x_i-z,m_s,s}\left(\omega-z,0,(\boldsymbol{\theta}\oplus z)^{(0_h)},0,\zeta^*,0\right)\Big]\bar{r}_s^Y(m_c,m_s)\Big)\bar{r}^{k,n}(f_s,k)\Big\} q^{k,n}(z) f^Y(y\,|\,t_s) r^{k,n}(f_s,k)\,dy\,dz \\
+\int_0^{i\tau}\sum_{s=1}^{n_1}\Big\{\Big[\psi(z,z,2) + G_{x_i-z,m_s,s}\left(\omega-z,t_s+z,\boldsymbol{\theta}\oplus z,f_s,\zeta^*,m_c\right)\Big] r^{k,n}(f_s,k) \\
+\Big[\psi(z,z,2) + G_{x_i-z,m_s,s}\left(\omega-z,t_s+z,(\boldsymbol{\theta}\oplus z)^{(0_h)},f_s,\zeta^*,m_c\right)\Big]\bar{r}^{k,\mathbf{n}}(f_s,k)\Big\} q^{k,n}(z)\mathcal{R}^Y(z\,|\,t_s) f^Z(z|t_s)\,dz \\
+\int_0^{i\tau}\Big[\Big(\psi(y,i\tau,0) + G_{x_{i-1},m_s,s}\left(i\tau,t_s+y,\boldsymbol{\theta}\oplus i\tau,f_s+1,\zeta^*,m_c+1\right)\Big) r_s^Y(m_c,m_s) + \Big(\psi(y,i\tau,1) + G_{x_{i-1},m_s,s}\left(i\tau,0,\boldsymbol{\theta}\oplus i\tau,0,\zeta^*,0\right)\Big)\bar{r}_j^Y(m_c,m_s)\Big] f^Y(y\,|\,t_s)\,dy\,\mathcal{R}^Z(\sigma|t_s) \\
+\Big(\psi(i\tau,i\tau,2) + G_{x_{i-1},m_s,s}\left(i\tau,t_s+i\tau,\boldsymbol{\theta}\oplus i\tau,f_s,\zeta^*,m_c\right)\Big)\mathcal{R}^Y(i\tau|t_s)\mathcal{R}^Z(i\tau|t_s). \qquad \text{(A.2)}
\end{aligned}
$$

Equation (A.2) is provided for a given inspection policy index a and has to be applied for all inspection policies. To get an idea about how the generator function is constructed, consider the first line in Equation (A.2). The placeholder function $\psi(y,z,0)$ and the recursive term $G_{x_i-z,m_s,s}\left(\omega-z,t_s+y,\boldsymbol{\theta}\oplus z,f_s+1,\zeta^*,m_c+1\right)$ contribute to the expected number of minimal repairs, if $r_s^Y(m_c,m_s)=1$, i.e. if the number of minimal repairs before replacement is lower than the optimal value. The other terms are responsible for the expected number of repairs, or the expected uptime.

The hazard rate of a hard subsystem consisting of all hard-type components in series configuration is given as $\lambda^H(z\,|\,\boldsymbol{\theta}) = \sum_{h=1}^{n_2}\lambda_h(\theta_h + z)$.

For the system with hard-type and soft-type components, Equations (A.1-A.2) can still be used, if function $r^{k,n}(f_s,k)$ is redefined as $r_h^Z\left(z,\zeta_h^*\right)$:

$$
r_h^Z\left(z,\zeta_h^*\right) = \begin{cases} 1, \textit{ if } z \le \zeta_h^* \\ 0, \textit{ otherwise} \end{cases},
$$

and $q^{k,n}(z)$ is redefined as $q^h(z)$, which is the probability that the failure of the hard subsystem at time $Z=z$ is due to the failure of hard-type component h:

$$
q^h(z) f^Z(z|\boldsymbol{\theta}) = \lambda_h(z\,|\,\boldsymbol{\theta}_h)\exp\left(-\int_0^Z \lambda^H(\chi\,|\,\theta)\,d\chi\right).
$$

References

1. Aven T, Dekker R. A useful framework for optimal replacement models. Reliability Engineering and System Safety 1997; 58(1): 61-67, https://doi.org/10.1016/S0951-8320(97)00055-0.
2. Babishin V, Taghipour S. Joint maintenance and inspection optimization of a k-out-of-n system. Proceedings of the Annual Reliability and Maintainability Symposium 2016: 523-528, http://dx.doi.org/10.1109/RAMS.2016.7448039.
3. Babishin V, Taghipour S. Joint optimal maintenance and inspection for a k-out-of-n system. International Journal of Advanced Manufacturing Technology 2016; 87(5): 1739-1749, https://doi.org/10.1007/s00170-016-8570-z.
4. Babishin V, Taghipour S. Optimal maintenance policy for multicomponent systems with periodic and opportunistic inspections and preventive replacements. Applied Mathematical Modelling 2016; 40(23-24): 10480-10505, http://dx.doi.org/10.1016/j.apm.2016.07.019.
5. Bjarnason E T S, Taghipour S. Periodic inspection frequency and inventory policies for a k-out-of-n system. IIE Transactions. 2016; 48(7): 638-650, https://doi.org/10.1080/0740817X.2015.1122253.
6. Bjarnason E T S, Taghipour S, Banjevic D. Joint optimal inspection and inventory for a k-out-of-n system. Reliability Engineering and System Safety 2014; 131: 203-215, https://doi.org/10.1016/j.ress.2014.06.018.
7. Castanier B, Grall A, Bérenguer C. A condition-based maintenance policy with non-periodic inspections for a two-unit series system. Reliability Engineering and System Safety 2005; 87: 109-120, https://doi.org/10.1016/j.ress.2004.04.013.
8. Cho D I, Parlar M. A survey of maintenance models for multi-unit systems. European Journal of Operational Research 1991; 51: 1-23, https://doi.org/10.1016/0377-2217(91)90141-H.
9. Cui L, Li H. Opportunistic maintenance for multi-component shock models. Mathematical Methods of Operations Research 2006; 63: 493-511, https://doi.org/10.1007/s00186-005-0058-9.
10. Dagpunar J S. A maintenance model with opportunities and interrupt replacement options. The Journal of the Operational Research Society 1996; 47(11): 1406-1409, https://doi.org/10.1057/jors.1996.176.
11. Dekker R, Wildeman R E, van der Duyn Schouten F A. A review of multicomponent maintenance models with economic dependence. Mathematical Methods of Operations Research 1997; 45: 411-435, https://doi.org/10.1007/BF01194788.
12. Flage R, Aven T. Optimal periodic condition inspection and replacement policy for a binary monotone system using a counting process approach. Proc. IMechE Part O: J. Risk and Reliability 2010; 225: 161-168.
13. Gao W, Zhang Z, Ji H, Zhou Y, Liu Q. Optimal quasi-periodic preventive maintenance policies for a repairable system with stochastic maintenance interval. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17(3): 389-397, https://doi.org/10.17531/ein.2015.3.9.
14. Golmakani H R, Makedi H. Optimal non-periodic inspection scheme for a multi-component repairable system using A* search algorithm. Computers & Industrial Engineering 2012; 63: 1038-1047, https://doi.org/10.1016/j.cie.2012.07.002.
15. Gunn E A, Diallo C. Optimal opportunistic indirect grouping of preventive replacements in multicomponent systems. Computers & Industrial Engineering 2015; 90: 281-291, https://doi.org/10.1016/j.cie.2015.09.013.
16. Hajipour Y, Taghipour S. Non-periodic inspection optimization of multi-component and k-out-of-n systems using genetic algorithm. Reliability Engineering and System Safety 2016; 156: 228-243, https://doi.org/10.1016/j.ress.2016.08.008.
17. Lapa C M F, Pereira C M N A, Frutuoso e Melo P F. Surveillance test policy optimization through genetic algorithms using non-periodic intervention frequencies and considering seasonal constraints. Reliability Engineering and System Safety 2003; 81: 103-109, https://doi.org/10.1016/S0951-8320(03)00085-1.
18. 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.
19. Lienhardt B, Hugues E, Bes C, Noll D. Failure-finding frequency for a repairable system subject to hidden failures. Journal of Aircraft 2008; 45(5): 1804-1809, https://doi.org/10.2514/1.31149.
20. Ozekici S. Optimal periodic replacement of multicomponent reliability systems. Operations Research 1988; 36(4): 542-552, https://doi.org/10.1287/opre.36.4.542.
21. Pandey M, Zuo M J, Moghaddass R. Selective maintenance scheduling over a finite planning horizon. Proc. IMechE Part O: J. Risk and Reliability 2016; 230(2): 162-177.
22. Park K S. Optimal number of minimal repairs before replacement. IEEE Transactions on Reliability 1979; R-28(2): 137-140.
23. Peng W, Liu Y, Zhang X, Huang H Z. Sequential preventive maintenance policies with consideration of random adjustment-reduction features. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17(2): 306-313, https://doi.org/10.17531/ein.2015.2.19.
24. Sheu S H, Li S H, Chang C C. A generalised maintenance policy with age-dependent minimal repair cost for a system subject to shocks under periodic overhaul. International Journal of Systems Science 2012; 43(6): 1007-1013, https://doi.org/10.1080/00207720802645220.
25. Sriram C, Haghani A. An optimization model for aircraft maintenance scheduling and re-assignment. Transportation Research Part A 2003; 37: 29-48, https://doi.org/10.1016/S0965-8564(02)00004-6.
26. Su B. An optimal inspection and diagnosis policy for a multi-mode system. Reliability Engineering and System Safety 2002; 76: 181-188, https://doi.org/10.1016/S0951-8320(02)00003-0.
27. Taghipour S. Optimal inspection model for a load-sharing redundant system. Proceedings of the Annual Reliability and Maintainability Symposium 2014; 198-202.
28. Taghipour S. Reliability and maintenance of medical devices. Toronto: LAP Lambert Academic Publishing, 2011.
29. Taghipour S, Banjevic D. Optimal inspection of a complex system subject to periodic and opportunistic inspections and preventive replacements. European Journal of Operational Research 2012; 220(3): 649-660, https://doi.org/10.1016/j.ejor.2012.02.002.
31. Taghipour S, Banjevic D. Optimum inspection interval for a system under periodic and opportunistic inspections. IIE Transactions 2012; 44: 932-948, https://doi.org/10.1080/0740817X.2011.618176.
32. Taghipour S, Banjevic D. Periodic inspection optimization models for a repairable system subject to hidden failures. IEEE Transactions on Reliability 2011; 60(1): 275-285, https://doi.org/10.1109/TR.2010.2103596.
33. Taghipour S, Kassaei M L. Periodic inspection optimization of a k-out-of-n load-sharing system. IEEE Transactions on Reliability 2015;

64(3): 1116-1127, https://doi.org/10.1109/TR.2015.2421819.
34. Vaurio J K. Optimization of test and maintenance intervals based on risk and cost. Reliability Engineering and System Safety 1995; 49(1): 23-36, https://doi.org/10.1016/0951-8320(95)00035-Z.
35. Wang H. A survey of maintenance policies of deteriorating systems. European Journal of Operational Research 2002; 139: 469-489, https://doi.org/10.1016/S0377-2217(01)00197-7.
36. Wang KH, Kuo C C. Cost and probabilistic analysis of series systems with mixed standby components. Applied Mathematical Modelling 2000; 24: 957-967, https://doi.org/10.1016/S0307-904X(00)00028-7.
37. Wang H, Pham H. Reliability and optimal maintenance. London: Springer, 2006.
38. Yun WY, Endharta AJ. A preventive replacement policy based on system critical condition. Proc. IMechE Part O: J. Risk and Reliability. 2016;230(1):93-100.
39. Zhang YL, Wu S. Reliability analysis for a k/n(F) system with repairable repair-equipment. Applied Mathematical Modelling 2009; 33: 3052-3067, https://doi.org/10.1016/j.apm.2008.10.022.
40. Zhao X, Fouladirad M, Bérenguer C. Residual-based inspection/replacement policy for a deteriorating system with Markovian covariates. Proceedings of the Industrial Engineering and Engineering Management (IEEM) Conference 2010; 1: 636-641, https://doi.org/10.1109/IEEM.2010.5674526.
41. Zhu W, Fouladirad M, Berenguer C. A reactive multi-component maintenance policy for offshore wind turbines. In: European Safety and Reliability Conference (ESREL) 2013; 811-817.
42. Zille V, Bérenguer C, Grall A, Despujols A. Modelling multicomponent systems to quantify reliability centred maintenance strategies. Proc. IMechE Part O: J. Risk and Reliability 2011; 225: 141-160.

Vladimir BABISHIN
Yassin HAJIPOUR
Sharareh TAGHIPOUR
Department of Mechanical and Industrial Engineering
Ryerson University
350 Victoria Street
Toronto, Ontario, M5B 2K3, Canada

E-Mails: vbabisin@ryerson.ca, yassin.hajipour@ryerson.ca, sharareh@ryerson.ca

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

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. Sławczo Denczew
The Main School of Fire Service, Warsaw, Poland

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. 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, 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.

All the scientific articles have received two positive reviews from independent reviewers.

Our 2016 Impact Factor is 1.145

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 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 zloty 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 hab, inż. Marek Hawryluk
Department of Mechanical Engineering
Wroclaw University of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
E-mail: marek.hawryluk@pwr.edu.pl

Prof. dr hab. inż. Zbigniew Gronostajski
Department of Mechanical Engineering
Wroclaw University of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
E-mail: zbigniew.gronostajski@pwr.edu.pl

Dr inż. Jacek Ziemba
Department of Mechanical Engineering
Wroclaw University of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
E-mail: jacek.ziemba@pwr.edu.pl

Dr inż. Łukasz Dworzak
Department of Mechanical Engineering
Wroclaw University of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
E-mail: lukasz.dworzak@pwr.edu.pl

Mgr inż. Paweł Jabłoński
Department of Mechanical Engineering
Wroclaw University of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
E-mail: pawel.jablonski@pwr.edu.pl

Mgr inż. Macin Rychlik
Kuźnia Jawor S.A.
Ul. Kuziennicza 4, 59-400 Jawor, Poland
E-mail: marcinrychlik@kuznia.com.pl

ANALIZA WPŁYWU WARUNKÓW SMAROWANIA NA ZUŻYCIE NARZĘDZI W PROCESACH KUCIA MATRYCOWEGO NA GORĄCO

Streszczenie:
Praca dotyczy problematyki smarowania w procesach kucia matrycowego na gorąco z uwzględnieniem trwałości narzędzi i oprzyrządowania kuźniczego. Przedstawiono badania literaturowe oraz własne autorów dotyczące wpływu zastosowania środków smarno-chłodzących, ilości dawki i kierunku jej podawania oraz innych czynników wpływających na warunki tribologiczne. Przeanalizowano także obecnie stosowane w przemyśle urządzenie i systemy smarowania. Na tej podstawie autorzy w oparciu o wiedzę i doświadczenie opracowali i zbudowali urządzenie smarujące. Opracowany system, zaimplementowany do przemysłowego procesu pozwala na dobór i zapewnienie optymalnych warunków tribologicznych w procesie poprzez sterowanie ilością i częstotliwością podawanej dawki środka smarnego. Może być ono alternatywą dla manualnej metody nanoszenia środka smarnego, zależnej od czynnika ludzkiego lub w pełni zautomatyzowanych, lecz drogich systemów smarowania. Uzyskane wyniki badań wskazują na potencjalne możliwości wprowadzenia na stałe do pracy zbudowanego urządzenia także do innych procesów kucia poprzez integrację z manipulatorem. Zaproponowane rozwiązanie zapewnia większą stabilność i powtarzalność warunków smarowania oraz pozytywnie wpływa na zwiększenie wydajności procesu wytwarzania, a tym samym znacząco obniża jednostkowe koszty produkcji odkuwek.

Słowa kluczowe: *kucie matrycowe na gorąco, system smarujący, zużycie, warunki tribologiczne*

1. Wstęp

Podczas kucia matrycowego na ciepło i gorąco narzędzia poddawane są bardzo dużym, cyklicznym obciążeniom cieplnym od 80 °C do 600 °C oraz mechanicznym, sięgającym nawet powyżej 1200 MPa [37]. Do głównych i najczęstszych mechanizmów niszczących zaliczyć można [13,26]: zużycie ścierne [4], pękanie mechaniczne, odkształcenie plastyczne [23] oraz zmęczenie cieplne [9,36] i cieplno-mechaniczne [7]. Przebieg obciążenia matryc oraz stempli ma cykliczny charakter. Stanowi ono połączenie obciążenia cieplnego i mechanicznego, które wynika z kontaktu i odkształcania gorącego materiału przedkuwki poprzez zimne narzędzie. Szczególny wpływ na trwałość matryc ma zmienne obciążenie cieplne. Jest ono główną przyczyną powstawania pęknięć zmęczeniowych oraz zmian własności fizyczno-mechanicznych wierzchniej warstwy narzędzi [3,30]. Jednocześnie obciążenie cieplne wpływa na zintensyfikowanie zużycia ściernego wywołanego wysokimi naciskami mechanicznymi, a to przekłada się na obniżenie czasu eksploatacji narzędzi i oprzyrządowania kuźniczego [25, 26].

W przypadku trwałości narzędzi kuźniczych kluczowe jest smarowanie. Standardowo podczas kucia matrycowego półswobodnego (spęczanie, spłaszczanie, itp.) nie stosuje się smarowania narzędzi ze względu na nieskomplikowane płynięcie odkształcanego materiału. Inaczej niż ma to miejsce w typowym kuciu matrycowym, w trakcie którego użycie środka smarno-chłodzącego jest wymagane ze względu na konieczność minimalizowania tarcia w celu dokładnego wypełnienia przez odkształcany materiał wykroju roboczego narzędzia. Zastosowanie środków smarno-chłodzących powoduje przede wszystkim zmniejszenie tarcia między materiałem kutym, a materiałem narzędzi, jak również izoluje materiał narzędzia (wykrój roboczy) od bezpośredniego kontaktu z gorącym materiałem kutym. W efekcie skutecznie obniża to temperaturę powierzchni matrycy i tym samym zmniejsza intensywność procesów odpuszczania, utleniania i erozji [13]. Wadą wynikającą ze stosowania środków smarno-chłodzących jest gwałtowane schłodzenie warstwy wierzchniej, co może przyspieszać proces zmęczenia cieplnego. Dodatkową funkcją środka smarnego jest zmniejszenie współczynnika tarcia po uwolnieniu kutego metalu z matrycy [1,2,4]. Na odpowiedni sposób smarowania i chłodzenia narzędzi kuźniczych wpływają nie tylko właściwości środka smarnego, lecz także sposób i kierunek jego podawania, ilość i częstotliwość ekspozycji dawki smarnej. Istotne jest także zapewnienie powtarzalnego i równomiernego rozprowadzania środka smarnego. Obecnie wciąż, w dużej liczbie kuźni matrycowych stosowane są często niezautomatyzowane, mało powtarzalne urządzenia smarujące, które obsługiwane są przez operatorów w sposób manualny. Z tego też powodu coraz powszechniejsze staje się zastosowanie mniej lub bardziej zautomatyzowanych urządzeń oraz systemów smarno-chłodzących, pozwalających na precyzyjne podawanie środka smarnego [12,18,22]. Odpowiednie smarowanie decyduje nie tylko o powstawaniu odkuwki bez wad, takich jak niewypełnienia, ale także korzystnie wpływa na zmniejszenie zużycia narzędzi [20]. Wraz z rozwojem automatyzacji, zaczęły powstawać zautomatyzowane urządzenia oraz systemy smarujące [5,11,17]. Rozwiązania takie są proste w użyciu i zostały z powodzeniem zaimplementowane w przodujących w automatyzacji produkcji kuźniach niemieckich. Dodatkowo dzięki wykorzystaniu manipulatorów, możliwe jest zbudowanie w pełni elastycznego sytemu pozwalającego na kontrolowanie wszystkich istotnych parametrów smarowania, jak pozycja dyszy, czas nakładania, proporcje substancji, itd. [10,18]. Dodatkowo systemy są zsynchronizowane z pracą agregatu kuźniczego, co przy sposobie podawania smaru, eliminuje subiektywny czynnik ludzki doświadczonego operatora [19]. Wiąże się to jednak z dosyć wysokimi kosztami inwestycji, które mogą się nie zbilansować w przypadku rozpatrywania jedynie zysku wynikającego z podniesienia jakości produktu oraz trwałości narzędzi.

Dlatego obecnie prowadzone są intensywne poszukiwania nowych rozwiązań dotyczących zarówno środków smarnych, dedykowanych do konkretnych aplikacji kuźniczych, jak i nowoczesnych zautomatyzowanych urządzeń smarujących zapewniających optymalne warunki tribologiczne, które w podwyższonych temperaturach są jeszcze bardziej krytyczne dla przemysłowych procesów kucia matrycowego na półgorąco i gorąco. Prowadzenie dalszych prac badawczo-rozwojowych w tym obszarze jest w pełni uzasadnione zarówno od strony naukowej, jak i finansowej, bowiem zagadnienie skutecznego smarowania stanowi nadal nierozwiązany problem i jest sporym wyzwaniem dla wielu ośrodków naukowych oraz firm przemysłowych.

Celem pracy jest analiza wyników badań dotyczących wpływu ilości, częstotliwości oraz sposobów podawania środka smarnego, a także zastosowania autorskiego systemu smarująco-chłodzącego na trwałość oprzyrządowania kuźniczego w procesach kucia matrycowego na gorąco.

2. Analiza stanu zagadnienia oraz badania dotyczące czynników wpływających na eksploatację narzędzi w procesach kucia matrycowego

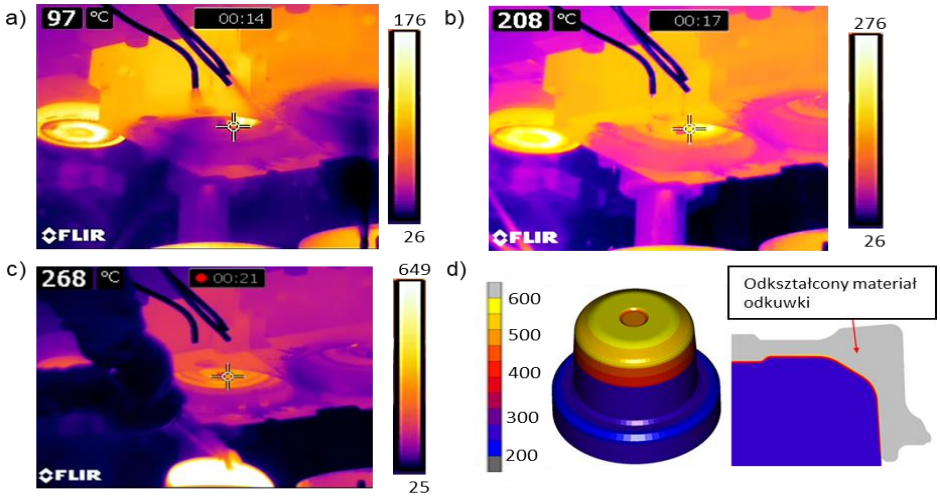
Należy podkreślić, że badaniom i analizie właściwości środków smarnych stosowanych w procesach kucia na gorąco, w dostępnej literaturze poświęcono wiele miejsca [1,11,12,21,24,33]. Ponadto nieustanny rozwój technologiczny sprawia, że poza ośrodkami naukowymi, coraz więcej badań oraz aplikacji przemysłowych, związanych ze środkami smarnymi, dedykowanych dla konkretnych procesów można znaleźć na bardzo konkurencyjnym rynku środków smarnych. Do wiodących producentów środków smarnych na świecie zaliczyć można takie firmy jak: Acheson, Fuchs, Henkel/Bechem, Houghton, Oelheld, itp. a także z Polski firma Naftochem.

Obok niepodważalnej roli środków smarnych, istotnymi czynnikami wpływającymi na smarowanie są odpowiedni sposób nanoszenia środka smarnego oraz technologia jego dozowania. Obecnie nadal najpopularniejszą metodą smarowania jest bezpośredni, ręczny natrysk środka smarnego

wykonywany przez kowala podczas procesu. Główną wadą tego typu smarowania jest nierównomierne rozprowadzanie środka smarnego, co powoduje zmienny rozkład temperatury matrycy i w efekcie szybsze jej miejscowe zużycie. Jedynie w przypadku pras transferowych stosowane są obecnie w przemyśle w pełni zautomatyzowane systemy i urządzenia smarujące, zsynchronizowane z pracą całego agregatu.

W zdecydowanej większości prac w literaturze przedmiotu najwięcej miejsca poświęca się systemom smarowania w procesach kucia na zimno [29], głównie do aluminium [31], bądź na zimno i ciepło [5, 6, 28]. A zdecydowanie mniej, jeżeli chodzi o systemy i urządzenia przeznaczone do procesów kucia na ciepło i gorąco [27,32,37]. Nawet, w obszernej, przeglądowej pracy Altana [1], znaleźć można jedynie skrótowe informacje na temat systemów smarowania, które sprowadzają się bardziej do doboru smarów, aniżeli urządzeń smarujących. Z kolei w pracy [28] znaleźć można informacje na temat urządzeń oraz systemów w jakie wyposażone są agregaty kuźnicze stosowane w japońskim przemyśle kuźniczym, lecz także w cytowanej pracy brak jest danych o systemach smarowania. Interesujące są także wyniki Europejskiego Projektu Badawczego „Brite-Euram" [30] dotyczące opracowania przyjaznych dla środowiska systemów do smarowania narzędzi w procesach kucia stali na ciepło. Przykładowo w pracy [2] w badaniach przeprowadzonych przez zespół niemiecko-japoński na podstawie eksperymentów w trakcie produkcji, opracowano wytyczne, m.in.: schemat doboru optymalnych systemów smarowania w celu zwiększenia trwałości matrycy. Z kolei w pozycji [20] przedstawiono wybrane czynniki wpływające na jakość smarowania oraz zużycie narzędzi. Na uwagę zasługuje praca [35], w której autorzy zaprezentowali możliwości modelowania kierunku i ilości dawki środka smarnego.

Jak wykazano we wstępie, w procesach kucia na gorąco panują ekstremalne warunki pracy, co powoduje, że są one jednymi z najtrudniejszych do realizacji procesów wytwarzania. Autorzy artykułu od wielu lat prowadzą liczne badania dotyczące wpływu warunków tribologicznych na eksploatację narzędzi kuźniczych [15,16]. Przykładowo na rys. 1 przedstawiono wyniki badań dotyczące wpływu podawania środka smarnego na zmiany temperatury na powierzchni narzędzi w trakcie cyklu kucia.



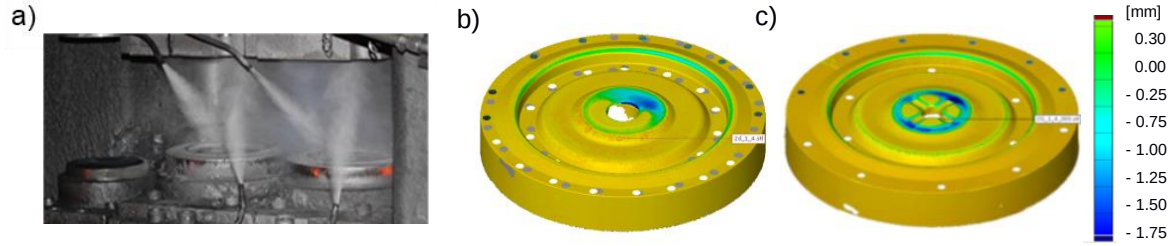
Rys. 1. Termogramy z rozkładami temperatur na górnych stemplach stosowanych w procesie kucia odkuwki typu tarcza: a) cykl smarowania narzędzi – temperatura w zaznaczonym punkcie wynosi 97 ^{0}C, b) 3 sekundy po zakończeniu smarowania – wzrost temperatury o ponad 100 ^{o}C, c) tuż po procesie kucia – około 1 sekundy po odkształceniu, d) pole temperatury na analizowanym stemplu w momencie kucia (kontaktu) wyznaczone z modelowania numerycznego

Jak można zaobserwować, na podstawie termogramów z kamery termowizyjnej (rys. 1a do rys. 1c), w ciągu zaledwie około 7 sekund procesu kucia, w momencie końcowej fazy smarowania trwającej około 2-3s (rys. 1a) temperatura na powierzchni narzędzia w wybranym punkcie wynosi około 100 ^{o}C (rys. 1a). Po kolejnych 3 sekundach temperatura wzrasta gwałtowanie do średniej „roboczej" temperatury pracy narzędzi w procesie kucia (rys. 1b). Tuż po kuciu – procesie odkształcenia – kolejne 4 sekundy – wskutek zamiany pracy odkształcenia na ciepło temperatura powierzchni narzędzi wzrosła o 60 ^{o}C (rys. 1c).

Prowadzone ponadto badania z wykorzystaniem modelowania numerycznego wykazały, że w momencie kontaktu kształtowanego materiału ze stemplem temperatura wzrasta do ponad 500 ^{o}C (rys. 1d). Obserwowane tak dynamiczne zmiany temperatury na powierzchniach roboczych narzędzi, przy nieodpowiednio dobranych parametrach smarowania mogą znacząco skrócić ich czas eksploatacji poprzez nadmierne cykliczne przechłodzenie, bądź nieukierunkowane i niezsynchronizowane z pracą agregatu

kuźniczego podawanie dawki środka smarnego. Wydaje się, że nawet najlepiej dobrany środek smarny nie wpłynie znacząco na istotne zmiany w rozkładach temperatur.

Na rys. 2a przedstawiono standardowy sposób smarowania stosowany w prasach korbowych, poprzez ręczne ustawienie dysz smarujących, zgodnie z uznaniem i doświadczeniem operatora.



Rys. 2a) Ręczny sposób dostarczania środka smarnego, b) zużycie narzędzia w przypadku nierównomiernego smarowania oraz c) zużycie narzędzia w przypadku równomiernego smarowania

Brak dostatecznej kontroli nad procesem smarowania oraz złe ustawienie dysz na początku procesu jest przyczyną nierównomiernego zużycia (rys. 2b i rys. 2c), przez co uszkodzenia następują intensywniej i znacząco skracają czas eksploatacji narzędzi oraz negatywnie wpływają na kształt i jakość odkuwki. Niezsynchronizowanie ekspozycji środka smarnego z pracą agregatu kuźniczego oraz przekładaniem odkuwki do kolejnych wykrojów (rys. 3a), a także nadmierne smarowanie może spowodować wady odkuwki oraz być przyczyną przedwczesnego zużywania narzędzi (rys. 3b).

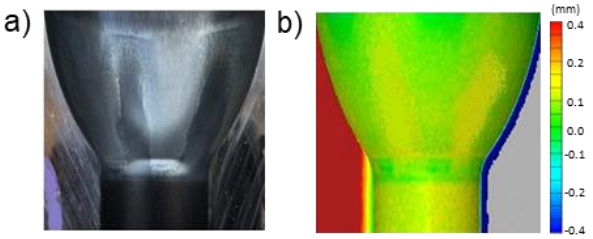
Rys. 3. Nieodpowiedni sposób smarowania: a) zła synchronizacja układu smarującego, b) zbyt intensywne smarowanie – brak odparowania cieczy z mieszaniny środka smarnego, c)pęknięcie w narożu wkładek matrycujących – efekt Rebindera

Niewłaściwie dobrana dawka środka chłodząco-smarującego może być przyczyną przedwczesnego pękania narzędzi oraz powstawania niewypełnień wykroju. Jest powodem zwiększenia wartości ciśnienia spowodowanego występowaniem kieszeni powietrznej w tym miejscu, przyspiesza powstawanie mikropęknięć oraz tzw. efektu Rebindera (rys. 3c). Rozwiązaniem przedstawionego problemu może być wyeliminowanie czynnika ludzkiego poprzez wprowadzenie precyzyjnego systemu dozowania i sekwencjonowania środka smarnego.

Wraz z rozwojem automatyzacji, zostały opracowane rozwiązania, które pozwalają na bardziej precyzyjne i dokładniejsze nanoszenie materiału smarnego. Obecnie można wyróżnić kilka typów rozwiązań technicznych układów do nanoszenia powłok smarnych [24]. Ze względu na wysoką wydajność oraz krótki czas nanoszenia środka smarnego, rozwiązania zautomatyzowane są zrealizowane w oparciu o natrysk bezpośredni. Jest on wykonywany poprzez system dysz, które rozpylają środek smarny na matrycy podczas zmiany materiału wsadowego. Budowa dyszy jest zależna od używanego środku smarnego oraz parametrów jak np. ciśnienie rozpylania, lepkość smaru. Zaletą takiej metody smarowania jest łatwość ukierunkowania strumienia środka smarnego oraz wytworzenie cienkiej warstwy izolacyjnej na powierzchni narzędzia. Ponadto zautomatyzowane układy zazwyczaj posiadają stację przygotowania środka smarnego lub jego mieszanki w celu utrzymania jednolitego składu w całej objętości. Wcześniejsze przygotowanie środka smarnego zmniejsza możliwość osadzania się cząstek stałych w przewodach oraz dyszach, co w efekcie powoduje wydłużenie czasu pomiędzy serwisami tych urządzeń. W celu uzyskania równomiernej powłoki natryskiwanej, dysze są ustawiane tak, aby odzwierciedlały krzywiznę narzędzia. Wadą użycia klasycznych dysz jest ich wysoki koszt, rozmiar oraz problem z zastosowaniem przy matrycach o dużych gabarytach.

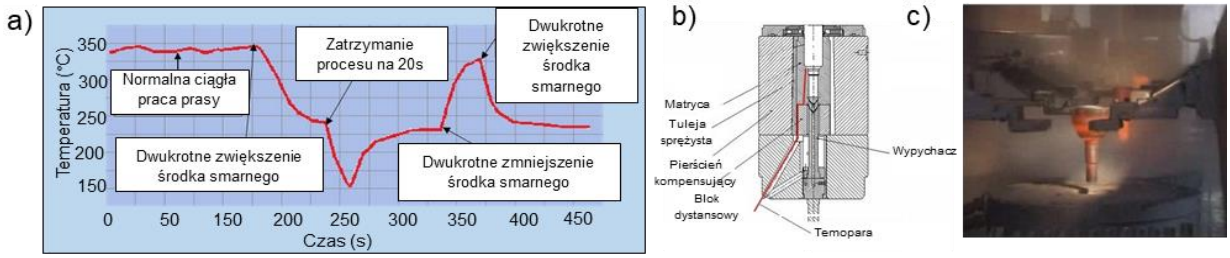
Sam proces automatyzacji procesu smarowania nie pozwala jednak na rozwiązanie analizowanego problemu. Przykładem takim może być zautomatyzowany sposób smarowania na prasie korbowej wyposażonej w transfer odkuwek, tak jak to ma miejsce na przykład w procesie kucia przegubu homokinetycznego. W procesie tym narzędzia smarowane są przez specjalnie skonstruowany wyrzutnik z dyszami. Sześć dysz rozmieszczonych równomiernie na obwodzie, co 60° dostarcza smar w momencie

wyrzutu odkuwki, gdy wyrzutnik znajduje się w górnym położeniu. Efektem tego jest pokrycie powierzchni wnętrza matrycy warstwą smaru z różną intensywnością zależnie od bliskości dyszy i nierównomierny stopień zużycia (wskutek adhezji) na obwodzie matrycy (rys. 4a).



Rys. 4. Nierównomiernie zużyta powierzchnia narzędzia: a) zdjęcie cyfrowe b) wyniki skanowania

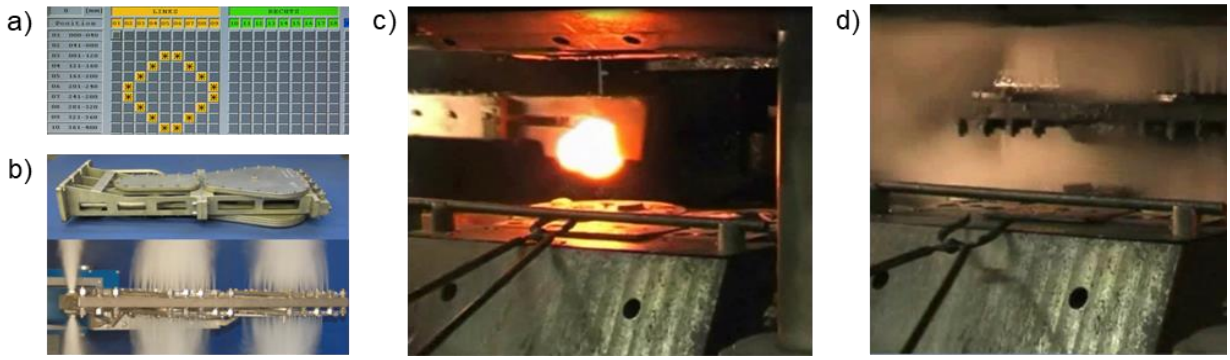
Na powierzchni narzędzi można zaobserwować wyraźną różnicę pomiędzy strefami zielonymi oraz żółtymi (rys. 4b), a ciemnymi, do których docierała mniejsza ilość płynu smarującego. W tym samym procesie w celu analizy wpływu ilości dawki smarnej na zmiany temperatury wewnątrz narzędzia (rys. 5) autorzy dokonali pomiaru tego parametru. Wykorzystano do tego termoparę typu K wprowadzoną do jednej z matryc. Pomiar wykazał, że dwukrotne zwiększenie wydatku środka smarnego powoduje spadek temperatury na powierzchni matrycy o około 100 °C. Ponadto analizując cały przebieg, można zaobserwować jak ważne jest optymalizowanie i kontrolowanie procesu smarowania dla całego okresu eksploatacji danego narzędzia a nie tylko pojedynczej operacji kucia.



Rys. 5a) Zmian temperatury w matrycy podczas kucia z zaznaczonymi najczęściej występującymi zakłóceniami w procesie, b) schemat zamontowania termopary w narzędziu, c) widok przenoszonych odkuwek na prasie transferowej

Podobne badania dotyczace temperatury na powierzchni narzędzi kuźniczych podczas cyklu kucia przedstawiono w pracy [8].

Obecnie coraz powszechniejszą metodą smarowania, zwłaszcza w nowoczesnych i zautomatyzowanych kuźniach staje się wykorzystanie ramion manipulatorów, które precyzyjnie oraz powtarzalnie wsuwają głowice smarujące pomiędzy narzędzia. Są to wysoko wydajne systemy, które ze względu na cenę i skomplikowaną budowę są częściej stosowane w kuciu na zimno [10]. Do wiodących producentów systemów smarujących opartych o manipulatory zaliczyć można: AED Automation, SMS Group, Renite, Spay. Natomiast obecnie chętnie wykorzystywanym, znacznie tańszym w stosunku do drogich manipulatorów, a przy tym wysoce zautomatyzowanym systemem smarującym przeznaczonym do kucia matrycowego na gorąco jest urządzenie opracowane przez firmę Jerko (rys. 6) [19]. System ten posiada specjalną, wsuwaną pomiędzy górne i dolne narzędzia płytę smarującą nanoszącą warstwę smaru. Rozwiązanie takie wykorzystane zostanie w Kuźni Jawor w ramach realizowanego projektu badawczego.



Rys. 6. Widok: a) panelu z systemu smarowania z układem aktywnych (wybranych dysz) firmy Jerko [19], b)widok płyty smarującej z dyszami, c) manipulator z chwytakiem wkładający przedkuwkę, d) wjazd płyty smarującej i uruchomienie smarowania

Dodatkowo urządzenie pozwala na programowy wybór dysz, które mają być aktywne podczas podawania środka smarującego, co dodatkowo zwiększa elastyczność systemu.

Jak wykazano powyżej, mimo niezaprzeczalnego, ciągłego rozwoju technologicznego oraz opracowywania nowych rozwiązań technicznych, obecnie nadal dla większości typowych agregatów kuźniczych stosowane są stare, często zawodne i mało powtarzalne urządzenia smarujące, które dodatkowo są obsługiwane przez operatorów, od których zależy jakość smarowania w procesie kucia. Na podstawie przedstawionych badań oraz analizy stanu literatury można zauważyć, że kluczowe zagadnienie w przypadku smarowania stanowi przede wszystkim sposób i kierunek podawania oraz ilość (objętość) oraz czas dozowania czynnika smarująco-chłodzącego. Czynniki te mają decydujące znaczenie, a przypadku ich niewłaściwego doboru powodują znaczne skrócenie czasu eksploatacji narzędzi, a nawet przedwczesne uszkodzenie oprzyrządowania. Odzwierciedla się to także na obniżonej jakości i dokładności wymiarowo-kształtowej odkuwek. Przeprowadzona analiza pozwala stwierdzić, że największy wpływ na zużywanie się narzędzi kuźniczych, pomijając aspekt doboru smaru, ma precyzyjne ustawienie objętości i częstotliwości ekspozycji dawki środka smarno-chłodzącego. Nieco mniejsze znaczenie ma kierunek podawania smaru względem powierzchni matrycy. Optymalnym kierunkiem natrysku jest kierunek normalny względem powierzchni narzędzia. Jednak ze względu na skomplikowane kształty matryc, jest to w większości przypadków niemożliwe do zrealizowania. Wraz ze wzrostem skomplikowania układu dozującego, tj. rozbudowania głowicy natryskowej, czy układu automatyki, znacząco wzrasta koszt układu smarowania. Obecnie istniejące komercyjne rozwiązania zostały zbudowane w oparciu o wiedzę i doświadczenie firm tworzących te systemy dla wybranych aplikacji przemysłowych. W literaturze brak jest informacji na temat konstrukcji i kryteriów prawidłowego doboru układu smarowania, zwłaszcza dla procesów kucia w podwyższonych temperaturach.

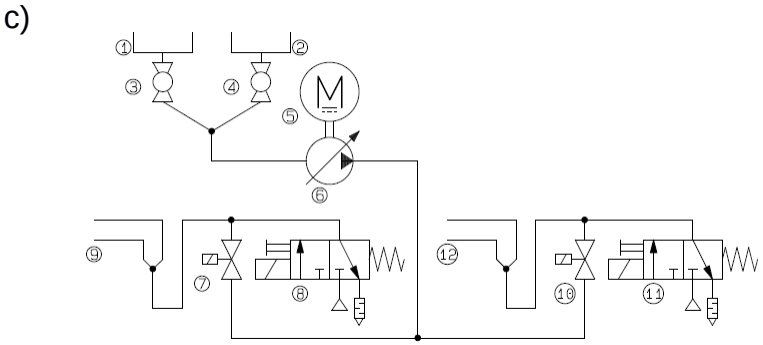
3. Koncepcja autorskiego systemu smarowania

W odpowiedzi na przedstawiony stan wiedzy, wyniki badań i zapotrzebowanie autorzy, na podstawie doświadczenia zbudowali prototypowy system smarująco-chłodzący dedykowany dla procesów kucia odkuwek typu rozwidlonego. Zbudowany układ dozująco – rozdzielający jest pierwszym etapem budowy kompleksowego, elastycznego układu smarowania. W kolejnym etapie zostanie on rozbudowany o manipulator i system dysz. Pozwoli to na prowadzenie dalszych badań i określenie optymalnej metody właściwego natryskiwania środka smarnego (rys. 7).

Jest to kolejne tego typu rozwiązanie opracowane przez autorów [16]. W dotychczas stosowanym urządzeniu smarującym w procesie kucia odkuwki rozwidlonej (widłaka), operator za pomocą zaworu „subiektywnie" regulował dawkę środka smarnego i czas jego ekspozycji poprzez dyszę smarującą. Powodowało to brak stabilności smarowania oraz zmienność warunków tribologicznych w procesie kucia. W porównaniu do poprzednich wersji, obecnie zbudowany system smarujący został częściowo zmodernizowany oraz użyto znormalizowanych elementów składowych, co pozwala na łatwe serwisowanie urządzenia w procesie eksploatacji przemysłowej oraz dostęp do powszechnie dostępnych części, co znacznie obniża koszty urządzenia. Urządzenie zostało wykonane i przetestowane na linii produkcyjnej w Kuźni Jawor, która jest głównym odbiorcą systemu. Przeprowadzone testy pozwoliły na potwierdzenie wysokiej sprawności i pozwoliły na poprawę procesu kucia na gorąco, co przekłada się na jakość i obniżenie kosztów produkcyjnych. W obecnym rozwiązaniu, szczególną uwagę zwrócono na cykl pracy urządzenia (kolejność sekwencji), ilość dawki oraz kierunek podawania środka smarnego.

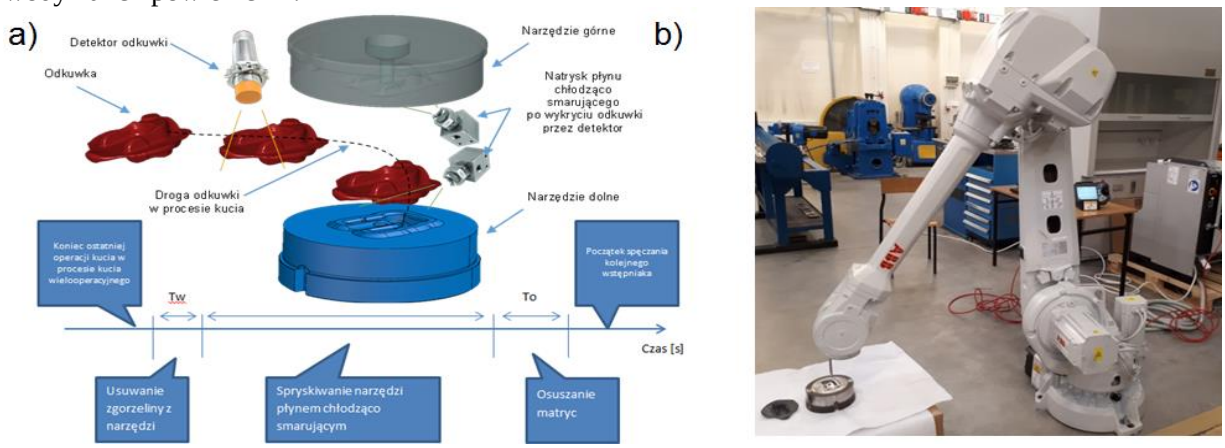
Opisywane rozwiązanie pozwala na precyzyjne i powtarzalne odmierzanie dozowanej dawki z jednoczesną możliwością zróżnicowania proporcji dla kolejnych dysz oraz możliwością swobodnego ustawienia sekwencji nadmuchu powietrza i nanoszenia środka smarnego. W tym celu w urządzeniu o budowie zaprezentowanej na schemacie (rys. 7c) wykorzystano pompę perystaltyczną (6) z silnikiem krokowym (5), która pozwala na precyzyjne odmierzanie ilości dozowanej cieczy smarującej (1).

a)  b) 



Rys.7. a) Zdjęcie urządzenia dozująco-smarującego pozwalające na regulację wydatku środka smarno-chłodzącego, b) widok podłączonego i pracującego urządzenia w trakcie przemysłowego procesu kucia, c) schemat układu pneumatyki

Dystrybucja dozowanej cieczy na dyszę górną (9) i dolną (12) realizowana jest poprzez dwa zawory membranowe (7 i 10) sterowane pneumatycznie (8 i 11). Po odmierzeniu ilości (objętości) cieczy rozdzielana jest dla każdej z dysz, tj. górnej i dolnej. Opracowane rozwiązanie pozwala na przepychanie środka smarnego do opracowanej przez autorów specjalnej dyszy wydmuchowej poprzez elastyczne przewody za pomocą sprężonego powietrza. Sprężone powietrze stosowane jest dodatkowo do rozpylania cieczy jest to tzw. powietrze dodmuchujące. Ustawienie kolejności podawania środka smarująco-chłodzącego, czasu przepychania i dodmuchu realizowane jest przez operatora za pomocą prostego w obsłudze panelu HMI. Odpowiednio dobrany czas wyrzutu i ilość środka smarującego sprzyja właściwemu rozpyleniu płynu, hamując procesy kumulacji grafitu w zagięciach matryc oraz pozostawania wody na ich powierzchni.



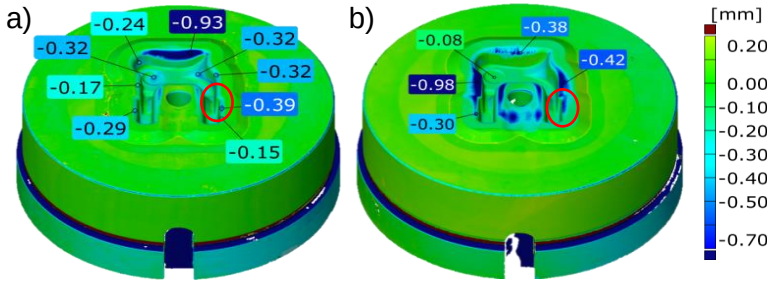
Rys.8. a) Idea działania systemu smarowania wraz z przykładowym schematem cyklu pracy, b) wykorzystanie ramienia manipulatora do ukierunkowanego podawania środka smarno-chłodzącego

Rys. 8a przedstawiono schemat działania urządzenia oraz przykładowy schemat cyklu pracy urządzenia. Regulacja zawartości fazy ciekłej w mieszaninie smarująco-chłodzącej polega na zmianie czasu działania pompy dozującej. Urządzenie posiada dodatkowy zbiornik z czystą wodą (oznaczony nr 2 na rys.7c) służący do czyszczenia i tym samym utrzymywania stałych parametrów smarowania. Wybór między zbiornikiem z cieczą smarującą, a czystą wodą odbywa się przy pomocy zaworów kulowych (oznaczony nr 3 i 4 na rys.7c). System jest ponadto wyposażony w mieszadło antysedymentacyjne, które pozwala na utrzymanie jednorodności zawiesiny grafitu w wodzie. Wstępnie prowadzone badania wykazały, wysoce zadowalające działanie urządzenia w warunkach przemysłowych. Kolejnym etapem rozwoju opracowanego urządzenia do podawania środka smarno-chłodzącego będzie integracja z manipulatorem (rys. 8b).

4. Porównanie i analiza wyników skanowania narzędzi kuźniczych eksploatowanych przy wykorzystaniu nowoopracowanego i dotychczas stosowanego systemu smarowania

Na rys. 9 przedstawiono przykładowe wyniki zbiorcze ze skanowania dolnych wkładek matrycowych stosowanych w operacji kucia wstępnie matrycującego odkuwki typu rozwidlonego po różnej liczbie odkuwek oraz po różnych warunkach tribologicznych (w wyniku zmiany systemu smarowania). Przed skanowaniem wkładki zostały oczyszczone z pozostałości smaru i zgorzeliny. Następnie zeskanowane

cyfrowe obrazy zostały przetworzone do postaci chmury punktów i zamienione na siatki trójkątów a następnie odniesione do modelu CAD.



Rys. 9. Wyniki skanowania wkładek ze stali UNIMAX stosowanych w procesie kucia na gorąco: a) po 16000 odkuwek z nowym systemem smarowania, b) po 16000 odkuwek ze starym systemem smarowania

Przedstawione przykładowe wyniki porównawcze skanów wkładek matrycowych (rys. 9) przetworzonych na cyfrowy obraz wskazują na to, że w przypadku zastosowania nowego urządzenia smarno-chłodzącego pozwalającego na precyzyjne ustawienia dawki smaru, zaobserwowano zmniejszenie zużycia narzędzi. Wstępne testy wykazały, że użycie dawki 10ml środka smarnego (mieszanina grafitu z wodą w stosunku 1:20) w czasie natrysku przez 2s, daje najlepsze rezultaty. Ponadto można także zaobserwować, że zapewnienie „lepszych" warunków smarnych narzędzia, w przypadku wkładki (rys. 9a) spowodowało pewne przesunięcia w zużyciu wybranych stref narzędzia, w stosunku do narzędzia ze starym systemem smarowania. Przykładowo dla narzędzia, przedstawionego na rys. 9b pojawiło się zużycie w miejscach, gdzie kształtowane są końcowe części ramion odkuwki, podczas gdy dla narzędzia pokazanego na rys. 9a (z nowym systemem smarującym) w tym obszarze brak jest śladów zużycia (zaznaczono czerwoną elipsą). Można to tłumaczyć tym, że w wyniku wprowadzenia do procesu nowoopracowanego systemu smarowania zmieniły się warunki tribologiczne, które spowodowały zmiany występowania mechanizmów niszczących w wybranych obszarach narzędzia.

W celu określenia historii zużywania się wybranych wkładek matrycowych (rys. 9) wykorzystano opracowaną przez autorów i częściowo rozbudowaną metodę skanowania odwrotnego, która jak wykazała praktyka przemysłowa stanowi użyteczne narzędzie pomiarowo-badawcze. Do tej pory metoda ta była aplikowana z powodzeniem do osiowosymetrycznych obiektów, o stosunkowo prostych kształtach i dużych wartościach ubytku materiału [14,15]. Opracowana metoda polega na pomiarze przy użyciu skanera, postępującego zużywania się wybranego narzędzia kuźniczego (rys.10a), w postaci jego ubytku materiału, a na podstawie zmian kształtu, cyklicznie pobieranych z procesu odkuwek, w postaci przyrostu materiału (rys.10b). W celu odtworzenia przebiegu zużycia narzędzia dokonuje się skanowania wybranych z serii produkcyjnej odkuwek (co 1000 szt.) z łącznej maksymalnej liczby wytworzonych odkuwek dla każdej z wytypowanych wkładek matrycowych. Do wyznaczenia wykresu opisującego zależności objętościowego zużycia w zależności od liczby wykonanych odkuwek w czasie procesu kucia konieczne jest wyliczenie zmiany objętości pomiarów kolejnych odkuwek do wyniku 100 odkuwki nominalnej, które zostały wyrównane z wykorzystaniem powierzchni referencyjnej.

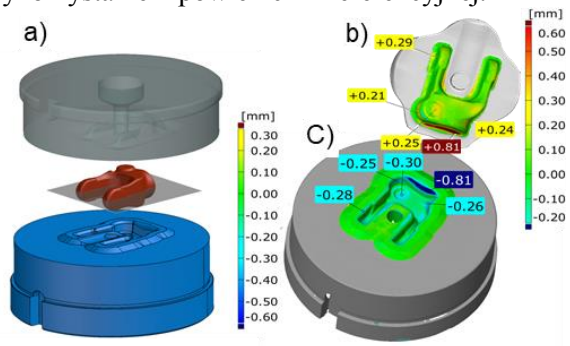
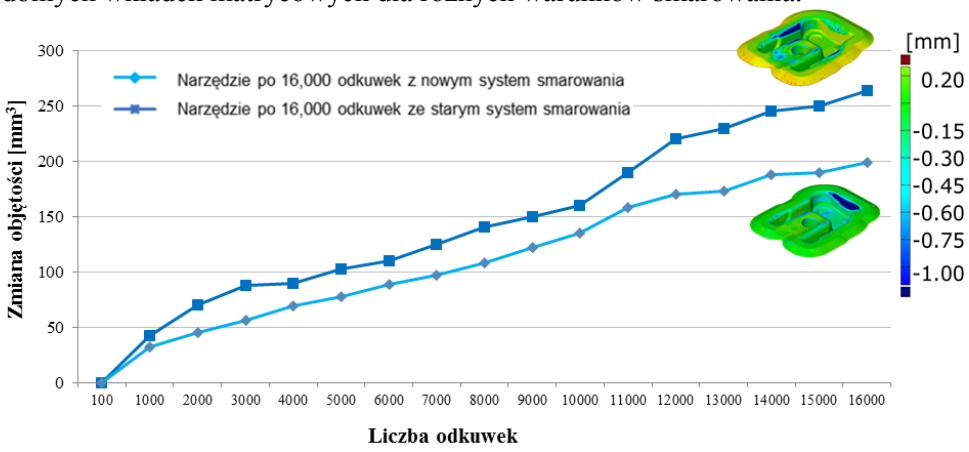


Fig. 10. Idea metody odwrotnej: a) komplet zestawu narzędzi kuźniczych dla drugiej operacji, b) porównanie wyników pomiarów ostatniej odkuwki oraz c) wkładki matrycowej po wykonaniu 16000 sztuk odkuwek, w postaci zmiany kształtu powierzchni

W ten sposób opracowuje się wykresy zmian objętości cyklicznie pobieranych w trakcie procesu i następnie skanowanych odkuwek. Pomiar zmian objętości ostatniej wytworzonej odkuwki dla danej wkładki matrycowej jest weryfikowany z pomiarem objętości narzędzia na koniec pracy. Na rys. 11

przedstawiono analizę zmian objętości przeprowadzoną w celu wyznaczenia historii zużywania się wybranych dolnych wkładek matrycowych dla różnych warunków smarowania.



Rys. 11. Porównanie krzywych zużycia (historii) wyznaczonych w oparciu o autorską metodę skanowania odwrotnego

Zaproponowane podejście wykorzystujące metodę skanowania odwrotnego 3d pozwala na globalny, ale zarazem pełny opis ubytku materiału, poprzez obserwowanie zmian objętości w trakcie procesu kucia, a także na szybką i miarodajną oraz praktyczną ocenę aktualnego stanu narzędzi kuźniczych. Ponadto uzyskane wyniki wskazują, że na podstawie przebiegu i kształtu (tendencji) krzywych zużycia można przewidywać wielkość ubytku materiału, a w konsekwencji sterować wielkością „otwarcia narzędzi" (grubość wypływki), a nawet wnioskować o momencie wycofania narzędzia z dalszej produkcji wskutek przekroczenia tolerancji wymiarowych.

Analizując wyznaczone przebiegi krzywych zużycia w aspekcie wpływu nowego systemu smarowania można wnioskować, że jego wprowadzenie do procesu zapewnia bardziej powtarzalne warunki tribologiczne i spowodowało zmniejszenie zużycia materiału zwłaszcza w końcowym okresie jego eksploatacji. Skłania to do przeprowadzenia bardziej zaawansowanych badań dotyczących zarówno optymalizacji dawki smaru, jak i kierunku jego podawania, częstotliwości oraz dodatkowych parametrów związanych z procesem smarowania.

5. Podsumowanie

W pracy przedstawiono analizę wpływu środków smarno-chłodzących, ilości dawki smaru, czasu podawania i kierunku oraz urządzeń smarujących na zużywanie się i eksploatację narzędzi kuźniczych. Jak wykazano, ilość środka smarnego oraz sposobu podaży stanowi nadal aktualny problem badawczy. Przeprowadzona analiza wykazała, że istotną kwestę stanowi optymalna dawka smaru oraz sposób jej podaży. Tylko poprzez kompleksową i długotrwałą analizę danego procesu kucia można dobrać optymalne warunki tribologiczne zapewniające stabilność i powtarzalność procesu wytwarzania. Dlatego obecnie duży nacisk kładziony jest na rozwój urządzeń smarno-chłodzących, gdyż to one mogą zapewnić powtarzalne warunki pracy narzędzi. W odróżnieniu od wciąż powszechnie stosowanego manualnego sytemu smarowania, na który wpływ ma subiektywny czynnik ludzki. Mimo, iż procesy kucia matrycowego w podwyższonych temperaturach są do siebie zbliżone, to jednak każdy wymaga indywidulanego podejścia. Powoduje to, że dla każdego procesu należy podchodzić oddzielnie i dobierać indywidualnie, jak i pozostałe parametry jego ekspozycji oraz sposób smarowania,. Przedstawiony w pracy, opracowany przez autorów system smarująco-chłodzący pozwala na dobór i zapewnienie optymalnych warunków tribologicznych w procesie, a także zbadanie wpływu ilości wielkości i częstotliwości dawki środka smarnego na zużywanie się narzędzia dla danego procesu. Wstępne wyniki badań wykazały przydatność tego urządzenia w procesie kucia matrycowego odkuwki typu rozwidlonego, w miejsce obecnie mało stabilnych i sterowanych manualnie przez kowala urządzeń smarujących. Uzasadnione są dalsze, bardziej zaawansowane prace nad doborem optymalnego wydatku środka smarnego oraz innych ustawień smarowania, co może się także wiązać z modernizacją opracowanego urządzenia. Prowadzone będą także dalsze badania związane zarówno z doborem środka smarno-chłodzącego, a także jego optymalnej temperatury, w celu zmniejszenia gradientu temperatur pomiędzy rdzeniem a powierzchnią roboczą narzędzi powodujących zmęczenie cieplne. W ramach dalszych prac planuje się wykorzystanie ramienia manipulatora do badań i testów nad doborem prawidłowego ukierunkowania i precyzją podawania środka smarnego do wykroju narzędzia.

Źródło finansowania: Badania zostały sfinansowane przez Narodowe Centrum Badań i Rozwoju (NCBiR); projekt POIG.01.03.01-02-063/12.

Literatura

1. Altan T, Ngaile G, Shen G. Cold and hot forging: fundamentals and applications. ASM International, Ohio 2005.
2. Altan T, Shirgaokar M. Advanced die materials and lubrication systems to reduce die wear in hot and warm forging. https://www.forging.org/uploaded/content/media/AltansPres.pdf
3. Anders P, Hogmark S, Bergström J. Simulation and evaluation of thermal fatigue cracking of hot work tool steels. International Journal of Fatigue 2004; (10): 1095-1107.
4. Archard J.F. Contact and rubbing of flat surfaces. Journal of Applied Physics 1953; (24): 981-988.
5. Bay N. New lubricant systems for cold and warm forging – advantages and limitations. In Liewald, M.: Proceed. 12th Int. Cold Forging Congr., Stuttgart, Germany 2011; 1(8).
6. Bay N. New Tribo-systems for Cold Forming of Steel, Stainless Steel and Aluminium Alloys. Proceedings of 46th International Cold Forging Group (ICFG) Plenary Meeting 2013.
7. Berti G.A, Monti M. Thermo-mechanical fatigue life assessment of hot forging die steel. Fatigue & Fracture of Engineering Materials & Structures 2005; 28 (11): 1025–1034.
8. Buchmayr B. Damage, Lifetime, and Repair of Forging Dies. BHM Berg- und Hüttenmännische Monatshefte 2017; 162 (3): 88–93.
9. Choi Ch, Groseclose A, Altan T. Estimation of plastic deformation and abrasive wear in warm forging dies. Journal of Materials Processing Technology 2012; 212 (8): 1742–1752.
10. Colin S.H. A review of automation in manufacturing illustrated by a case study on mixed-mode hot forging. Manufacturing Review 2014; (1)15, DOI: 10.1051/mfreview/2014012
11. Daouben E, E, et a., Effects of lubricant and lubrication parameters on friction during hot steel forging. Internationalof Journal Material Forming 2008; 1: 1223–1226.
12. Deacon R.F, Goodman J.K. Spreading Behavior of water based graphite Lubricants on Hot Die Surfaces. CIRP 2006; (55)1: 299-302.
13. Gronostajski Z, et al. The failure mechanisms of hot forging dies. Materials Science and Engineering, A Structural Materials: Properties, Microstructure and Processing 2016; 657: 147-160.
14. Gronostajski Z, Hawryluk M, et al. The application of the reverse 3D scanning method to evaluate the wear of forging tools divided on two selected areas. International Journal of Automotive Technology 2017; 18 (4): 653−662.
15. Gronostajski Z, Hawryluk M, Kaszuba M, Ziemba J. Application of a measuring arm with an integrated laser scanner in the analysis of the shape changes of forging instrumentation during production. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18(2): 194–200. http://dx.doi.org/10.17531/ein.2016.2.6.
16. Hawryluk M, et. al. Systems of supervision and analysis of industrial forging processes. Eksploatacja i Niezawodność - Maintenance and Reliability 2016; 18 (3): 315-324.
17. Hirschvogel M, Doelen H.V. Some applications of cold and warm forging. Journal of Materials Processing Technology 1992; 35 (7): 343-356.
18. https://www.aed-automation.com/en/products/solutions-forging.html, Aed Automation. Solutions forging. 2017. (accessed 17.11.02).
19. https://www.jerko-kempen.de, Jerko. 2017. (accessed 17.11.02).
20. Huskonen W.D. Trends in Die Lubrication. Forging Magazine 2004; (10): 24-26.
21. Isogawa S, Kimura A, Tozawa Y. Proposal of an evaluating method on lubrication. CIRP Annals - Manufacturing Technology 1992; 41: 263-266, 10.1016/S0007-8506(07)61200-1.
22. Iwama T, Morimoto Y. Die life and lubrication in warm forging. Journal of Materials Processing Technology 1997; 71: 43-48.
23. Kima D.H, Leeb H.C, Kimc B.M, Kimd K.H. Estimation of die service life against plastic deformation and wear during hot forging processes. Journal of Materials Processing Technology 2005; 166: 372–380.
24. Kumar U. et. al. Hot forging lubricants. International Journal of Mechanical Engineering and Robotics 2014; 3 (4): 155-163.
25. Lange K, Cser L, Geiger M, Kals J.A.G. Tool life and tool quality in bulk metal forming. Proceedings of the Institution of Mechanical Engineers. Part B: Journal of Engineering Manufacture November 1993; 207: 223-239.
26. Lavtar L, Muhic T, Kugler G, Tercelj M. Analysis of the main types of damage on a pair of industrial dies for hot forging car steering mechanisms. Engineering Failure Analysis 2011; 18 (10): 1143-1152.
27. Manji J. Die Lubricants, Forging 1994: 39–44.
28. Nagahama T, Enomae S. Cold- and warm-forging press developments and applications. Journal of Materials Processing Technology 1992; 35(3–4): 415-427.
29. Ngaile G, Saiki H, Ruan L, Marumo Y. A tribo-testing method for high performance cold forging lubricants. Wear 2007; 262: 684-692.
30. Persson A, Hogmarkb S, Bergstroma J. Thermal fatigue cracking of surface engineered hot work tool steels. Surface & Coatings Technology 2005; 191: 216– 227.
31. Sagisaka Y, Ishibashi I, Nakamura T, Sasaoka E, Hayakawa K. Evaluation of Environmentally Friendly Lubricant for Aluminium Alloy Cold Forging. Steel Res. Int., Special Edition, Wiley-VHC Verlag, Weinheim, 2011; 245-248.
32. Schey J. Tribology in Metalworking: Lubrication. Friction, and Wear American Society for Metals, USA 1983.
33. Sheljaskow S. Current level of development of warm forging technology. Journal of Materials Processing Technology 1994; 46 (7): 3-18.
34. Sheljaskow S. Tool lubricating systems in warm forging. Journal of Materials Processing Technology 2001; 113 (1–3): 16-21.
35. Soltani M, Pola A, La Vecchia G.M, Modigell M. Numerical method for modelling spray quenching of cylindrical forgings. La Metallurgia Italiana 2015; (7)8 : 33-40.
36. Starlinga C, Brancob J. Thermal fatigue of hot work tool steel with hard coatings. Thin Solid Films1997; 308(309): 436–442.

37. Taylan A, Gracious N, Gangshu S. Cold and hot forging fundamentals and application. ASM International. Asm metals handbook 2005; 14: 337-338.

Dr hab. inż. Andrzej Puchalski, prof. UTH[1]

Dr hab. inż. Marcin Ślęzak, prof. ITS[2]

Dr hab. inż. Iwona Komorska, prof. UTH[1]

Mgr inż. Piotr Wiśniowski[2]

[1] Uniwersytet Technologiczno-Humanistyczny w Radomiu
ul. Chrobrego 45, 26-600 Radom, Polska
[2] Instytut Transportu Samochodowego
ul. Jagiellońska 80, 03-301 Warszawa, Polska

E-mail: andrzej.puchalski@uthrad.pl
E-mail: marcin.slezak@its.waw.pl
E-mail: iwona.komorska@uthrad.pl
E-mail: piotr.wisniowski@its.waw.pl

Multifraktalna analiza eksploatacyjnego profilu prędkości pojazdu w zastosowaniu do testów jezdnych

Multifractal analysis vehicle's in-use speed profile for application in driving cycles

Słowa kluczowe: *analiza multifraktalna, testy jezdne, rzeczywiste warunki drogowe*

Streszczenie: Sygnały czasowe rejestrowane przez system diagnostyki pokładowej OBD i opisujące sposób ruchu pojazdu w rzeczywistych warunkach drogowych, wykazują niestacjonarność i nieliniowość oraz statystyczną wieloskalowość. W praktyce oznacza to, że analiza zarejestrowanych szeregów czasowych wymaga modelowania zjawisk nieliniowych. Celem niniejszej pracy było zbadanie charakteru profilu prędkości pojazdów w rzeczywistych warunkach drogowych metodą analizy multifraktalnej. Szereg badań wskazuje, że stosowane przez wiele lat testy jezdne nie były reprezentatywne dla rzeczywistych warunków eksploatacyjnych pojazdów. Zarówno dla nowego cyklu jezdnego WLTC, światowej zharmonizowanej procedury badań pojazdów lekkich jak i w rzeczywistych warunkach drogowych jazdy miejskiej na trasie pomiarowej, będącej przedmiotem badań doświadczalnych uzyskano potwierdzenie silnych własności multifraktalnych rejestrowanych szeregów czasowych prędkości pojazdu.

1. Wprowadzenie

Badania użytkowych właściwości samochodowych układów napędowych znajdują zastosowanie w fazie opracowywania koncepcji (podstawowe badania poznawcze), w fazie projektowania (w ramach badań prototypowych), w fazie produkcji (podczas kontroli jakości), na etapie dopuszczania do ruchu (badania homologacyjne) oraz w fazie eksploatacji (badania diagnostyczne). Cel ten jest realizowany w różnego rodzaju testach hamownianych, które różnią się narzucanymi warunkami pracy. W prezentowanej pracy podjęto problem analizy eksploatacyjnego profilu prędkości pojazdu w zastosowaniu do testów jezdnych. Cykl jezdny pojazdów (VDC- Vehicle Driving Cycle) to szereg czasowy danych reprezentujących prędkość

pojazdu, który ma odzwierciedlać ruch pojazdu w rzeczywistych warunkach drogowych i służyć ocenie ekonomicznej i emisyjnej pojazdu lub silnika.

Wieloletnie badania doprowadziły do kilkuset praktycznych cyklów jezdnych w różnych krajach i regionach. Większość stosowanych obecnie testów służących do wyznaczania emisji z pojazdów samochodowych jest opracowanych zgodnie z zasadą wiernej symulacji w dziedzinie czasu. Najbardziej popularne to testy takie jak FTP-75 (Federal Test Procedure), NEDC (New European Driving Cycle), czy JC08 (Japanese cycle). NEDC jest używany w Europie, cykl FTP 75 jest używany w Stanach Zjednoczonych a JC08 w Japonii. NEDC to cykl syntetyczny teoretycznego profilu jazdy, podczas gdy pozostałe dwa pochodzą z rzeczywistych danych dotyczących użytkowania [3, 9].

Bardziej miarodajne wyniki spalania w trakcie symulacji warunków eksploatacji pojazdu na hamowni, ma zapewnić test WLTC (Worlwide harmonized Light duty Test Cycle) będący rezultatem światowej zharmonizowanej procedury badań pojazdów lekkich. Cykl bazuje na rzeczywistych profilach jazdy pochodzących z badań statystycznych [18].

Każda metodyka tworzenia testów jezdnych wymaga analizy natury rejestrowanych sygnałów eksploatacyjnych prędkości pojazdu. Większość rzeczywistych sygnałów czasowych, w tym także szeregi czasowe rejestrowane przez system diagnostyki pokładowej OBD i opisujące sposób ruchu pojazdu w rzeczywistych warunkach drogowych, wykazuje niestacjonarność i nieliniowość oraz statystyczną wieloskalowość. W praktyce oznacza to, że analiza zarejestrowanych szeregów czasowych wymaga stosowania metod analizy nieliniowej [1, 7]. Metody formalizmu multifraktalnego, wykorzystujące lokalne miary o charakterze potęgowym, takie jak wykładniki Holdera regularności sygnału czy wskaźniki probabilistyczne, są dobrym sposobem modelowania dynamiki takich układów [5, 8] .

Wykładnik osobliwości Holdera wyznaczony w każdym punkcie, gdzie funkcja jest zdefiniowana, odzwierciedla poziom fluktuacji amplitudy w sąsiedztwie tego punktu. Czynnik skalujący funkcji rozdziału miary probabilistycznej bazującej na wartościach amplitudy sygnału, pozwala na segmentację uwzględniającą poziom entropii. Rozkład statystyczny reprezentowany przez histogram wykładników osobliwości stanowi graficzny obraz multifraktalności sygnału. Istnieją dwa sposoby wyznaczania widma multifraktalnego wykładników osobliwości. Metoda bezpośrednia polega na aproksymacji histogramu dla wyznaczonych wykładników Holdera. Metoda pośrednia, w której za wymiar fraktalny przyjmuje się zwykle wymiar pojemnościowy krzywej będącej wykresem rozważanego sygnału, bazuje na algorytmie analizy fluktuacji detrendowanych szeregu czasowego [6].

W artykule przedstawiono następujące zagadnienia. Rozdział 2 sygnalizuje ideę zastosowanej metody analizy multifraktalnej. Wyniki tej analizy w odniesieniu do cyklu jezdnego WLTC zharmonizowanej procedury badań pojazdów lekkich przedstawiono w rozdziale 3. Przeprowadzone badania w rzeczywistych warunkach drogowych jazdy miejskiej omówiono w rozdziale 4. Rozdział 5 stanowi podsumowanie.

2. Formalizm multifraktalny a miara regularności szeregu czasowego

Formalizm multifraktalny znalazł szereg zastosowań w analizie rzeczywistych sygnałów, a w szczególności w eksploatacyjnych badaniach diagnostycznych pojazdów samochodowych [4,11-17]. Prowadzone badania wskazują na różne procedury identyfikacji multifraktalnego charakteru rzeczywistych szeregów czasowych.

W niniejszym rozdziale zaprezentowano podejście bazujące na wykładnikach osobliwości sygnałów i widmie multifraktalnym. Punktowym wykładnikiem osobliwości Holdera funkcji $f(x)$ w punkcie x_0 nazywamy liczbę h zdefiniowaną jako kres górny wszystkich wykładników spełniających, dla pewnego C>0, warunek:

$$|f(x) - P_n(x - x_0)| \leq C|x - x_0|^h \quad (1)$$

gdzie $P_n(x - x_0)$ jest wielomianem stopnia $n < h$. Z zależności tej wynika, że wykładnik $h > 1$ opisuje regularność funkcji dokładniej niż jej kolejne pochodne. Jeżeli reprezentacją sygnału jest szereg czasowy $f_i, i = 1, 2, \dots, N$, to:

$$f_{i+\Delta i} - f_i \sim |\Delta i|^{h(i)} \quad (2)$$

gdzie $h(i)$ jest wykładnikiem osobliwości szeregu w punkcie i.

Zbiór wymiarów fraktalnych każdego podzbioru elementów szeregu czasowego f_i o tym samym wykładniku Holdera h tworzy multifraktalne widmo osobliwości. Do wyznaczenia widma wykładników Holdera sygnałów prędkości pojazdu rejestrowanych przez system OBD wykorzystano multifraktalną wersję analizy fluktuacji detrendowanych (MF-DFA) bazującą na eliminacji trendu z badanego szeregu czasowego. Procedura prowadzi do wyznaczenia miary w postaci funkcji fluktuacyjnej rzędu q, wykazującej charakter potęgowy:

$$F_q(s) \sim s^{H(q)} \quad (3)$$

gdzie $H(q)$ jest uogólnionym wykładnikiem Hursta, a parametr q pozwala zdekomponować miarę ze względu na jej wartość. Przekształcenie Legendre'a uogólnionego wykładnika wielkoskalowego $\tau(q) = qH(q) - 1$, pozwala na uzyskanie widma multifraktalnego:

$$f(h) = qh - \tau(q) \quad (4)$$

gdzie $h = \frac{d}{dq}\tau(q)$, to wykładnik osobliwości Holdera.

Widmo multifraktalne stanowi histogram wykładników odzwierciedlających poziom fluktuacji amplitudy sygnału. Ponadto paralelizm formalizmu multifraktalnego i termodynamiki statystycznej wskazuje, że wykładnik osobliwości Holdera i widmo multifraktalne można interpretować odpowiednio jako energię i entropię badanego procesu.

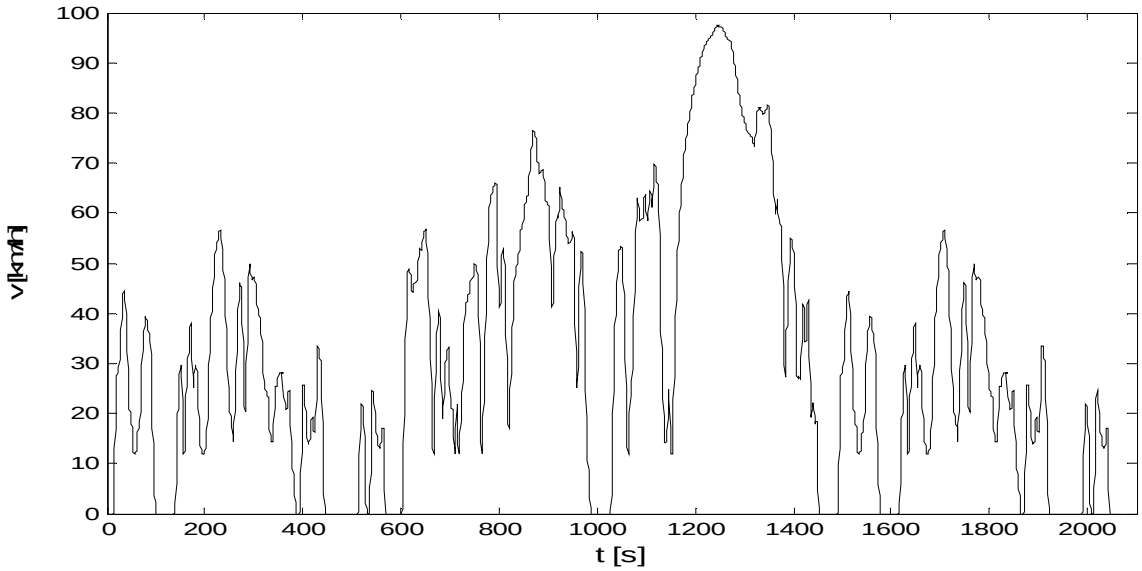
Opis własności dynamicznych układu na podstawie widma multifraktalnego szeregu czasowego umożliwiają:

- poziom multifraktalności $\Delta = h_{max} - h_{min}$ określony przez osobliwości o największej i najmniejszej fluktuacji szeregu czasowego (obserwowanego sygnału h_{min} i h_{max} ,
- osobliwość o największym wymiarze , czyli najczęściej spotykanej osobliwości szeregu czasowego $\{h_0 : f(h_{0)} = \max f(h)\}$,
- rozpiętość wymiarów podzbiorów osobliwości $\Delta f = f(h_{max}) - f(h_{min})$.

3. Badania symulacyjne testu WLTC

Szereg badań potwierdziło, że stosowane przez wiele lat testy jezdne nie są reprezentatywne dla rzeczywistych warunków eksploatacyjnych pojazdów. W związku z tym emisja i zużycie paliwa w pojazdach są niedoszacowane. Dążąc do bardziej dynamicznego zharmonizowanego cyklu testowego, opracowano nowy ogólnoświatowy cykl jezdny WLTC (Worlwide harmonized Light duty Test Cycle) . Syntezy dokonano na podstawie danych dotyczących parametrów ruchu w Europie, Indiach, Japonii, Korei i USA, uwzględniając

sytuacje związane z jazdą w ruchu miejskim, pozamiejskim i na autostradach. W porównaniu do NEDC, cykl testowania jest dłuższy, znacznie bardziej dynamiczny, ma dużo więcej cykli przyspieszania i hamowania, krótsze postoje oraz wyższą wartość prędkości średniej i maksymalnej. Ze względu na wprowadzone zmiany, cykl jezdny WLTC zapewni znacznie dokładniejsze warunki do obliczania zużycia paliwa i emisji spalin [10]. Jego pierwsze zastosowanie ma dotyczyć modeli pojazdów wprowadzone na rynek po raz pierwszy od września 2017 roku. Cykl jezdny WLTC został podzielony na cztery części, odpowiadające różnym prędkościom jazdy: niskiej, średniej, wysokiej i bardzo wysokiej. Jeżeli prędkość maksymalna nie przekracza 135km/godz., to część bardzo wysokiej prędkości należy zastąpić częścią prędkości niskiej. Przebieg czasowy cyklu WLTC, dla samochodów klasy 3b , ze stosunkiem mocy do masy PWr > 34, pokazano na rys.1, a jego podstawowe parametry w tabeli 1. Dynamika jazdy jest w przypadku tych pojazdów zdeterminowana przez zachowania kierowcy i intensywność ruchu drogowego, a nie przez parametry techniczne samochodu.



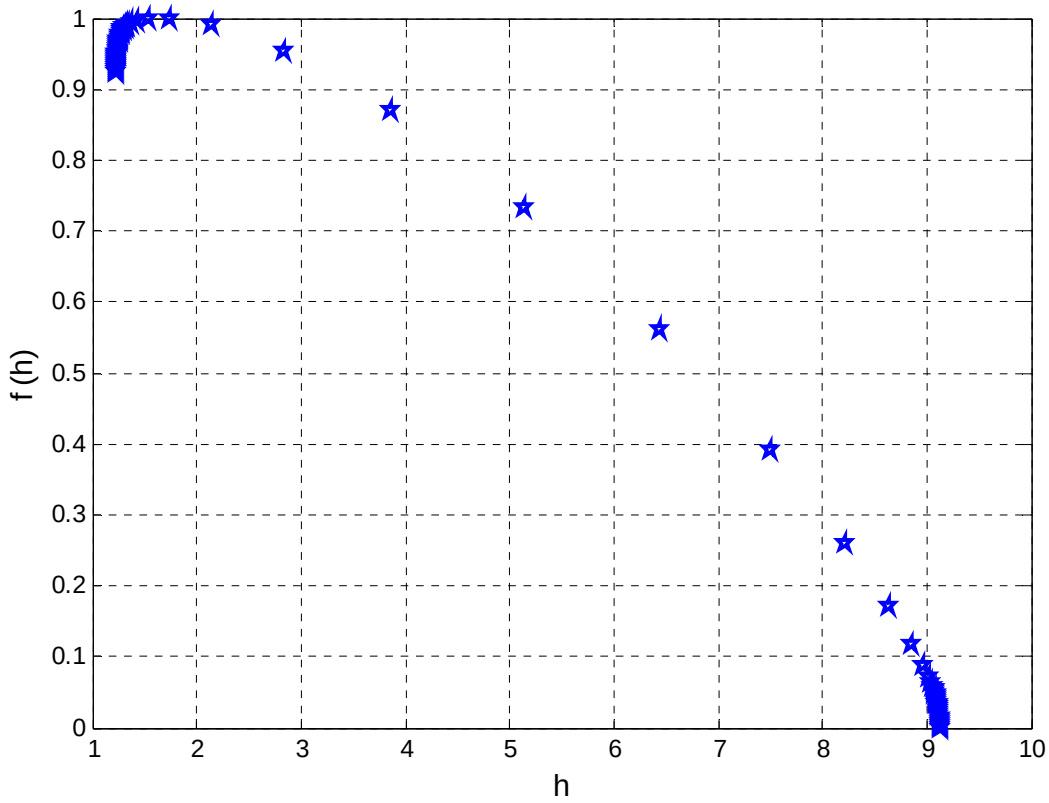
Rys.1. Cykl jezdny WLTC

Phase	Duration	Stop duration	Distance	p_stop	v_max	a_min	a_max
	s	s	m		km/h	m/s^2	m/s^2
Low	589	156	3095	26,5%	56,5	-1,47	1,47
Medium	433	48	4756	11,1%	76,6	-1,49	1,57
High	455	31	7162	6,6%	97,4	-1,49	1,58
Extra-High	323	7	8254	2,2%	131,3	-1,21	1,03
Total	1800	242	23266				

Tabela 1. Podstawowe parametry testu jezdnego WLTC dla klasy 3b

Widma osobliwości dla badanego testu jezdnego WLTC pokazane na rys.2, potwierdzają jego multifraktalną naturę. Poziom multifraktalności wynosi : $\Delta = 7{,}92$,

natomiast rozpiętość wymiarów segmentów o najwyższym i najniższym prawdopodobieństwie $\Delta f = 0{,}92$. Najczęściej spotykane wykładniki osobliwości dotyczą tych fragmentów zarejestrowanych szeregów czasowych, które opisuje największa zmienność. Najmniejsze prawdopodobieństwo rejestracji wykazują zaś okresy jazdy o najwyższej regularności. Większość punktów koncentruje się wokół wymiarów odpowiadających osobliwościom o największej i najmniejszej fluktuacji szeregu czasowego prędkości pojazdu.



Rys.2. Widmo osobliwości dla testu jezdnego WLTC

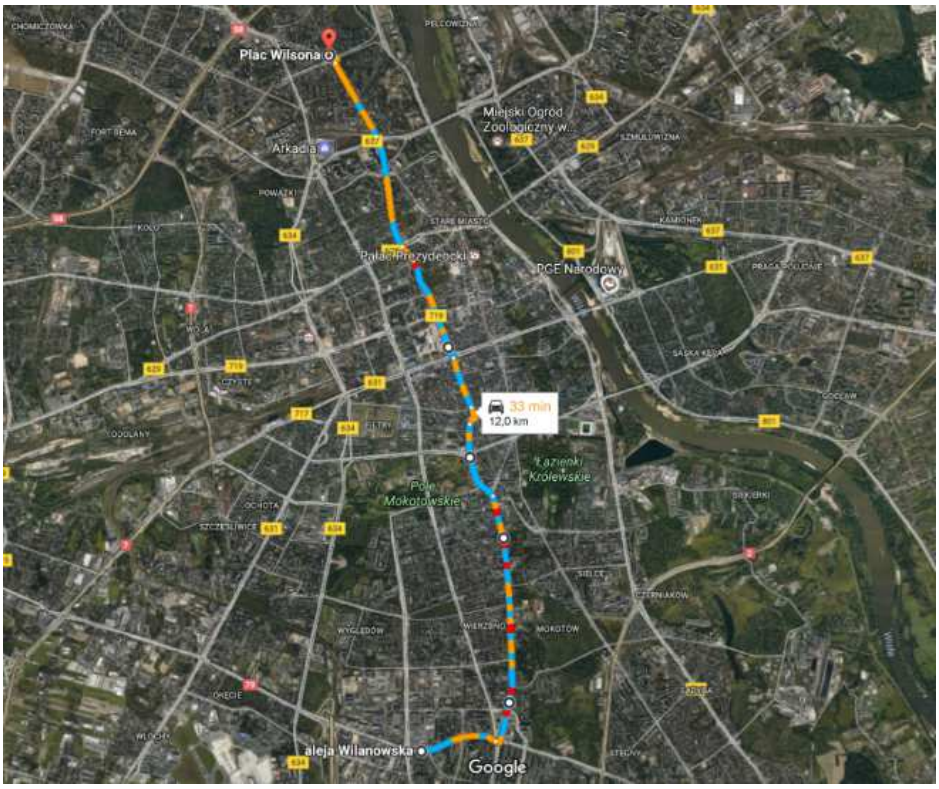
4. Badania i analiza danych doświadczalnych

W artykule przedstawiono wyniki badań ruchu pojazdów w rzeczywistych warunkach drogowych reprezentowanych przez jazdę miejską w dużej aglomeracji, w godzinach 9-13 przez pięć dni roboczych. Przeprowadzona analiza bazowała na szeregach czasowych prędkości pojazdu o okresie próbkowania równym 1s.

Wykorzystane urządzenie pomiarowe, widoczne na rys. 3 posiada możliwość próbkowania prędkości co 1 sekundę. Zakres prędkości możliwych do rejestracji to 0 -255 km/h, z rozdzielczością 1 km/h a maksymalny możliwy do zapisania dystans badania to 16 000 km. Urządzenie posiada możliwość rejestracji także szeregu innych parametrów pracy, jak obroty silnika, przyspieszenie pojazdu, ciśnienie w kolektorze dolotowym, czy położenie przepustnicy i umożliwia komunikację ze sterownikiem silnika z wykorzystaniem następujących protokołów łączności: J1850-41.6, J1850-10.4, ISO9141, KWP2000 (ISO 14230), CAN (Control Area Network ISO 11898).

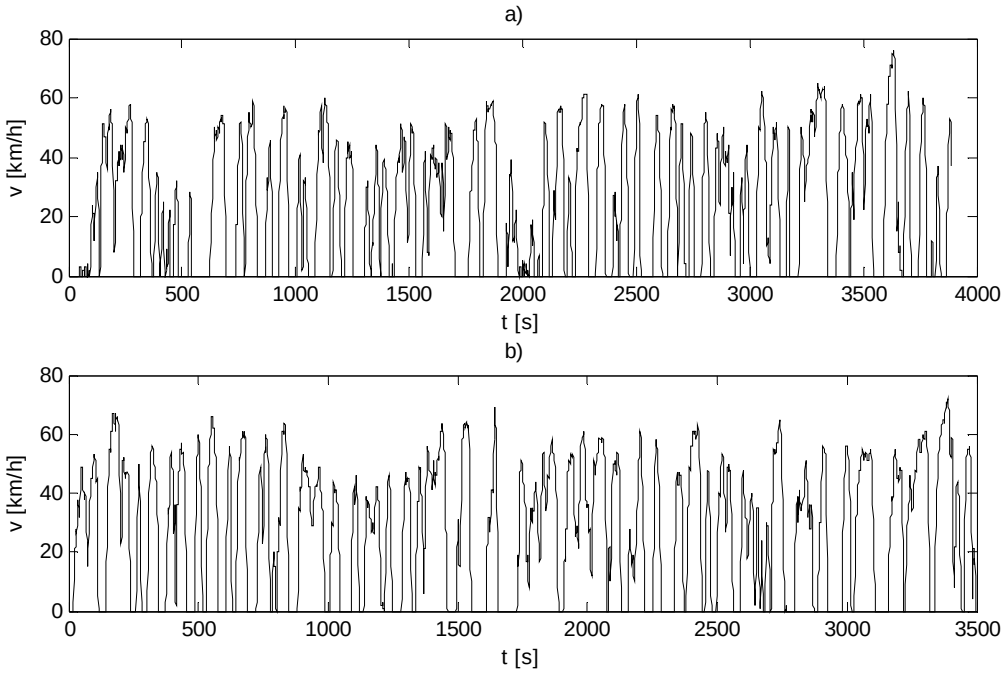
Rys. 3. Urządzenie rejestrujące CarChip Pro w gnieździe OBD pojazdu

Trasa pomiarowa obejmowała odcinek o długości 12 km i przebiegała pomiędzy Placem Wilsona a Galerią Mokotów w Warszawie. Mapę prezentującą trasę pomiarową, opracowaną z wykorzystaniem map Google, przedstawiono na rys. 4.



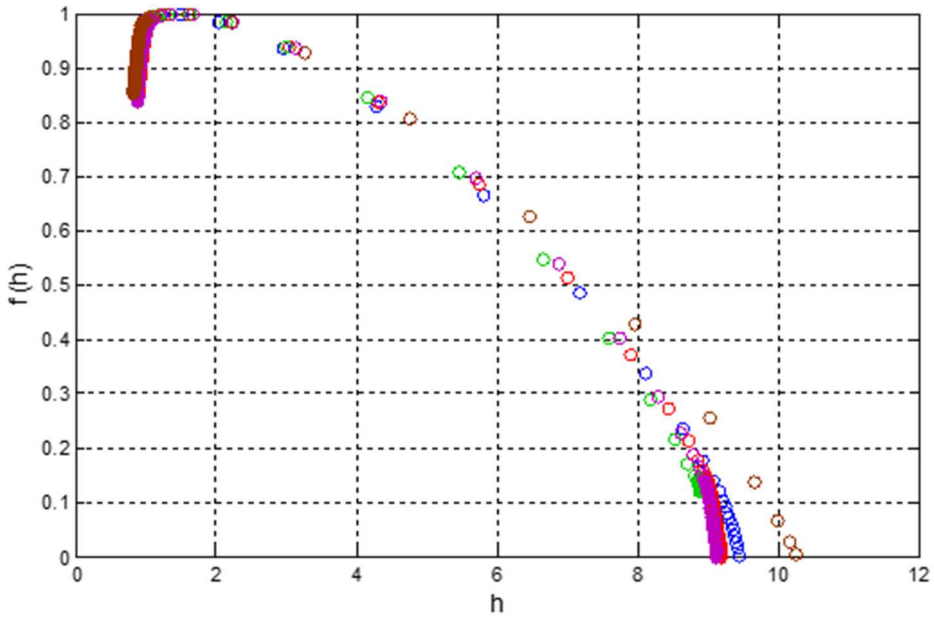
Rys. 4. Mapa prezentująca trasę pomiarową

Przykładowe przebiegi prędkości rejestrowane w ciągu jednego dnia dla jazdy w obie strony pokazano na rys. 5. Przeprowadzona analiza bazowała na szeregach czasowych prędkości pojazdu o okresie próbkowania równym 1s.

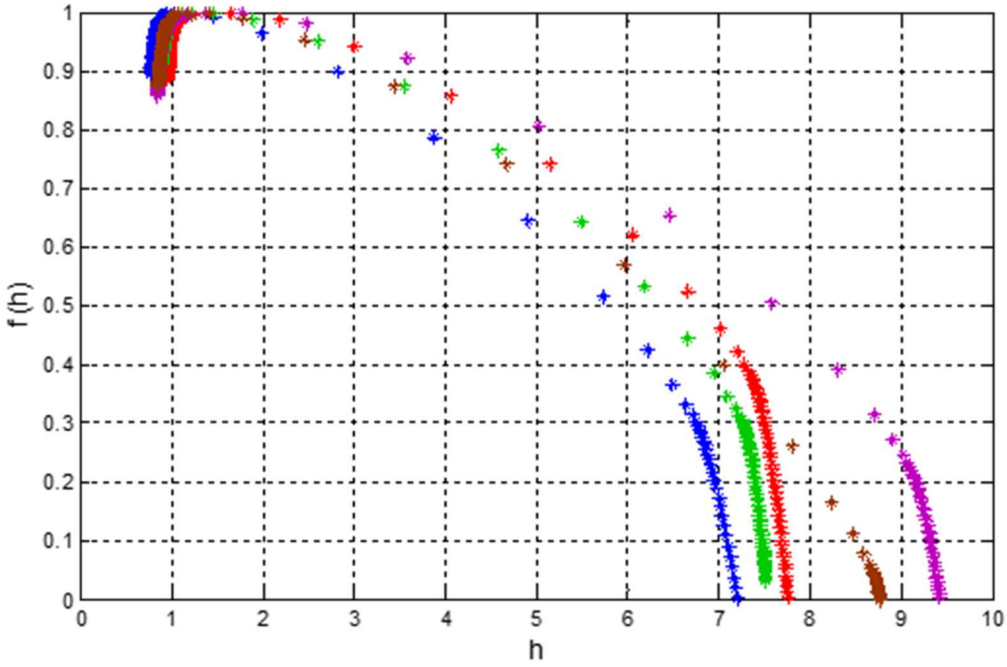


Rys. 5. Przykładowe przebiegi prędkości pojazdu

Widma multifraktalne pokazane na rys.6 - rys.7 są odzwierciedleniem dekompozycji rzeczywistego szeregu czasowego prędkości pojazdu za względu na częstość występowania fragmentów o określonej dynamice zmienności.



Rys. 6. Widma multifraktalne sygnałów prędkości pojazdu przy jeździe z punktu A do B



Rys.7. Widma multifraktalne sygnałów prędkości pojazdu przy jeździe z punktu B do A

Wszystkie spectra mają podobne własności w ramach przejazdu daną trasą w kolejne pięć dni roboczych. Wykazują zbliżone wartości poziomów multifraktalności, osobliwości o największym wymiarze i rozpiętości wymiarów podzbiorów osobliwości.

Segmenty zarejestrowanych sygnałów prędkości wykazujące największe fluktuacje osiągają wymiary $f(h) > 0.8$. Najmniejsze prawdopodobieństwo obecności wykazują okresy jazdy o regularności opisanej wykładnikiem osobliwości $h > 7$. Poziom multifraktalności Δ w zależności od dnia tygodnia, osiąga wartości w zakresie od 7 do 9.

4. Podsumowanie

Przeprowadzona analiza potwierdziła multifraktalny charakter fluktuacji prędkości pojazdu zarówno w ramach cyklu jezdnego WLTC światowej zharmonizowanej procedury badań pojazdów lekkich jak i w rzeczywistych warunkach drogowych jazdy miejskiej na tracie pomiarowej, będącej przedmiotem badań doświadczalnych w Warszawie. Widma multifraktalne charakteryzuje podobny kształt oraz zbliżone wartości poziomu multifraktalności. Niesymetryczny przebieg widm wskazuje na własności analogiczne dla multifraktali uzyskiwanych w wyniku realizacji uogólnionego procesu dwumianowej kaskady multiplikatywnej [2]. Uzyskane rezultaty wskazują na możliwość wykorzystania procesu kaskady multiplikatywnej do syntezy rzeczywistych warunków eksploatacyjnych oraz krótkookresowego prognozowania stanu ruchu drogowego, co stanowi krytyczny aspekt w rozwoju inteligentnych systemów transportowych. Krok ten wymaga prowadzenia dalszych badań obejmujących jazdę w warunkach pozamiejskich oraz na autostradzie.

Literatura

1. Batko W, Dąbrowski Z, Kiciński J. Nonlinear Effects In Technical Diagnostics. Radom, Poland: Institute for Sustainable Technologies, 2008.
2. Cheng Q. Generalized binomial multiplicative cascade processes and asymmetrical multifractal distributions. Nonlinear Processes in Geophysics 2014; 21: 477-487.
3. Chłopek Z. Synteza testów jezdnych zgodnie z kryteriami podobieństwa charakterystyk częstotliwościowych. Eksploatacja i Niezawodność – Maintenance and Reliability 2016; 18(4): 572-577.
4. Dai M, Zhang Ch, Zhang D. Multifractal and singularity analysis of highway volume data. Physica A 2014; 407: 332-340.
5. Kantelhardt JW. Fractal and Multifractal Time Series. Mathematics of Complexity and Dynamical Systems Robert A Meyers (Ed.). New York: Springer 2011.
6. Kantelhardt JW, Zschiegner SA, Koscielny-Bunde E, Havlin S, Bunde A, Stanley HE. Multifractal detrended fluctuation analysis of nonstationary time series. Physica A 2002; 316: 87-114.
7. Kantz H, Schreiber T. Nonlinear time series analysis. Cambridge: University Press, 2004.
8. Loutridis SJ. An algorithm for the characterization of time-series based on local regularity. Physica A 2007; 381: 383–398.
9. Merkisz J, Lijewski P, Fuc P, Siedlecki M, Ziółkowski A. Development of the methodology of exhaust emissions measurement under RDE (Real Driving Emissions) conditions for non-road mobile machinery (NRMM) vehicles. KONMOT 2016; Materials Science and Engineering 148: 1-11.
10. Mock P, Kühlwein J, Tietge U, Franco V, Bandivadekar A, German J. The WLTP: How a new test procedure for cars will affect fuel consumption values in the EU. Working paper 2014-9. International Council on Clean Transportation, 2014.
11. Puchalski A, Komorska I. Stable Distributions and Fractal Diagnostic Models of Vibration Signals of Rotating Systems. Advances in Condition Monitoring of Machinery in Non-Stationary Operations. Applied Condition Monitoring A.Timofiejczuk et al. (eds.) Springer International Publishing AG 2018; 9: 91-101.
12. Puchalski A. Multiscale analysis of vibration signals in engine valve system. Journal of Vibroegineering 2015; 17(7): 3586-3593.
13. Puchalski A. Techniki budowy nieliniowych modeli diagnostycznych z wykorzystaniem szeregów czasowych drgań mechanicznych. Przegląd Mechaniczny 2016; 10: 33-36.
14. Puchalski A, Komorska I. Multifractal Nature of Diesel Engine Rattle Noise in Vehicle. Archives of Acoustics 2017; 42(3): 469-474.
15. Shang P, Lu Y, Kamae S. The application of Holder exponent to traffic congestion warning. Physica A 2006; 370: 769–776.
16. Shang P, Lu Y, Kamae S. Detecting long-range correlation of traffic time series with multifractal detrended fluctuation analysis. Chaos, Solitions and Fractals 2008; 36: 82-90.
17. Shang P, Shen J. Multi-fractal analysis of highway traffic data Chinese Chinese Physics 2007; 16(2): 365-373.
18. Sileghem L, Bosteels D, May J, Favre C, Verhelst S. Analysis of vehicle emission measurements on the new WLTC, the NEDC and the CADC. Transportation Research Part D 2014; 32: 70–85.

dr inż. Witold Luty
Wydział Transportu
Politechnika Warszawska
Koszykowa 75, 00-662 Warszawa, Polska
E-mail: wluty@wt.pw.edu.pl

Symulacyjna analiza wpływu masy pojazdu na drogę zatrzymania

Słowa kluczowe: *bezpieczeństwo samochodu, droga zatrzymania, badania ogumienia*

Streszczenie: W pracy przedstawiono wyniki badań eksperymentalnych ogumienia pojazdu ciężarowego w warunkach dynamicznego hamowania. Na przykładzie wyników pomiaru pokazano, że zwiększenie obciążenia normalnego koła skutkuje wzrostem czasu narastania wzdłużnej reakcji stycznej oraz spadkiem wartości współczynnika przyczepności opony do podłoża (przylgowej oraz poślizgowej). Przedstawiono wyniki symulacji procesu hamowania awaryjnego pojazdu ciężarowego, którego masa zmienia się znacząco. Wyniki wykazały ze zwiększenie masy pojazdu może istotnie wydłużyć jego drogę zatrzymania w warunkach hamowania awaryjnego.

1. Wprowadzenie

Przyczepność kół jezdnych do podłoża może decydować zachowaniu pojazdu w ekstremalnych warunkach hamowania lub jazdy po łuku drogi na granicy przyczepności kół jezdnych do podłoża [13,14,15,17,3]. W procesie hamowania, długość drogi zatrzymania pojazdu można wyrazić jako [2,12,16]:

$$s_z = v_0 \left(t_{rk} + \mathrm{t_{rs}} + \frac{t_n}{2} \right) + \frac{v_0^2}{2a_h} \tag{1}$$

gdzie:

- a_h – średnie opóźnienie hamowania
- v_0 – prędkość początkowa pojazdu
- t_{rk} – czas reakcji kierowcy
- t_{rs} – czas reakcji układu hamulcowego
- t_n – czas narastania siły/opóźnienia hamowania

W warunkach hamowania awaryjnego, wartość opóźnienia hamowania a_h jest ograniczona siłą przyczepności kół jezdnych pojazdu do podłoża. W klasycznym ujęciu siła przyczepności każdego z kół jezdnych pojazdu zależy współczynnika przyczepności [1,17,3]. Zatem, w przypadku hamowania na poziomym podłożu, przy założeniu stałej wartości współczynnika przyczepności μ_2 kół jezdnych do podłoża (w stanie zablokowania kół), wyrażenie (1) przybiera następującą postać [2]:

$$s_z = v_0 \left(\mathrm{t_r} + \frac{t_n}{2} \right) + \frac{{v_0}^2}{2\mu_2 g} \tag{2}$$

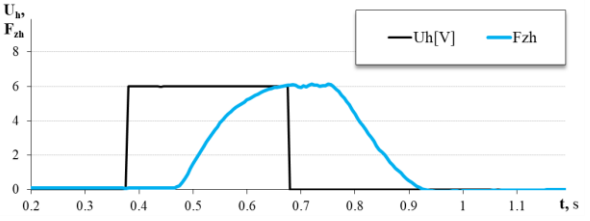
W przedstawionym wyrażeniu masa pojazdu nie ma udziału, co sugeruje, że nie wpływa na jego drogę zatrzymania podczas hamowana awaryjnego. Jednak doświadczenia własne oraz dane literaturowe pokazują, że wartość współczynnika przyczepności opony do podłoża może się zmniejszać wraz ze wzrostem wartości obciążenia normalnego koła [9,8,11], a wzrost masy pojazdu może wydłużyć drogę hamowania awaryjnego [18].

Zmiany masy (ciężaru) samochodów osobowych na ogół nie są znaczne, ale masa współczesnych samochodów ciężarowych może się zmieniać znacznie bo nawet 3 krotnie, od masy własnej do masy odpowiadającej Dopuszczalnej Masie Całkowitej (DMC).

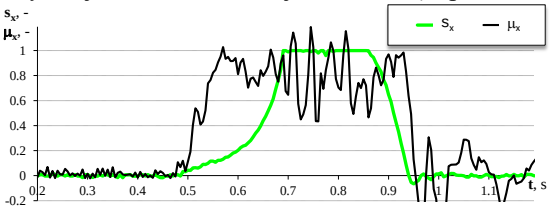
Wyniki własnych badań eksperymentalnych ogumienia pokazują, że wartość obciążenia normalnego koła zmienia przebieg zmian wartości wielkości fizycznych, które charakteryzują przebieg procesu hamowania. Pokazano to na rysunku 1. Przebieg całego procesu dynamicznego hamowania koła został opisany w innych publikacjach [7,8,6]. Na przykładowych wynikach pomiaru widać, że mimo zadania skokowego sygnału sterowania otwarciem zaworu pneumatycznego hamulca U_h, siła zacisku klocków hamulcowych na tarczy hamulcowej F_{zh} narasta z określonym przesunięciem w czasie, a jednocześnie z określonym, ograniczonym tempem, aż do osiągnięcia wartości maksymalnej. Z podobnym przesunięciem w czasie oraz z właściwym sobie tempem narastają: poślizg wzdłużny koła s_x oraz siła reakcji wzdłużnej F_x (zwana dalej reakcją wzdłużną) przenoszona przez koło. Przebieg tej siły pokazano na rysunku 1.b w postaci siły jednostkowej ($\mu_x = \frac{F_x}{F_z}$).

Przy określonej konstrukcji zacisku hamulcowego, wartość maksymalna siły zacisku klocków hamulcowych na tarczy hamulcowej F_{zh}, która determinuje wartość momentu hamowania koła M_h, jest ograniczona wartością ciśnienia powietrza dostarczonego do siłownika.

a) napięcie sterowania zaworem hamulca U_h oraz siła zacisku klocków hamulcowych F_{zh}



b) poślizg wzdłużny s_x i jednostkowa reakcja wzdłużna μ_x przenoszona przez koło



Rys. 1. Przykładowy zestaw wyników pomiaru wielkości fizycznych charakteryzujących proces dynamicznego hamowania koła samochodu ciężarowego średniej ładowności w warunkach laboratoryjnych (F_z=15000N, prędkość początkowa toczenia koła v_0=60km/h)

Zatem zmiana wartości obciążenia normalnego koła F_z nie powinna prowadzić do zmiany przebiegu oraz wartości maksymalnej siły zacisku hamulcowego F_{zh}, a tym samym momentu hamowania koła M_h. Jednak dynamika hamowania koła zmienia się. Widać to na rysunku 2.

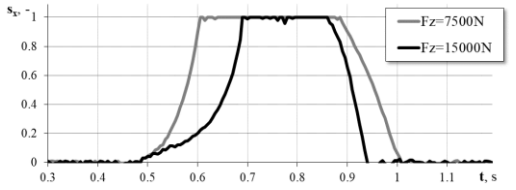
Wyniki pomiaru pokazują, że w wzrost obciążenia normalnego F_z, podczas dynamicznego hamowania koła, powoduje:

- wydłużenie czasu spadku prędkości kątowej koła, aż do jego zablokowania (Rys. 2a,b),
- wydłużenie czasu narastania wartości reakcji wzdłużnej F_x przenoszonej przez koło w wyniku hamowania do osiągnięcia wartości siły przyczepności w stanie zablokowania koła (Rys. 2b),
- spadek wartości współczynnika przyczepności przylgowej μ_1 oraz poślizgowej μ_2 koła do podłoża (Rys. 2c).

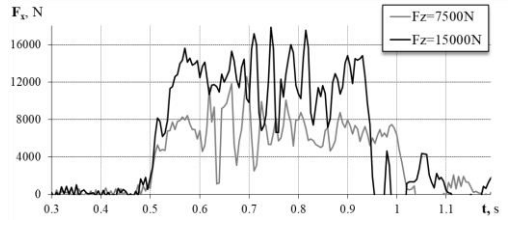
Na podstawie przedstawionych wyników badań laboratoryjnych koła ogumionego oraz w odniesieniu do równania (2) można stwierdzić, że wzrost masy pojazdu powodujący bezpośredni wzrost wartości obciążenia normalnego każdego z kół jezdnych, może spowodować wydłużenie jego drogi zatrzymania w procesie hamowania awaryjnego poprzez:

– wydłużenie czasu narastania siły hamowania, do osiągnięcia wartości odpowiadającej sile przyczepności kół jezdnych do podłoża,
– spadek wartości współczynnika przyczepności koła do podłoża.

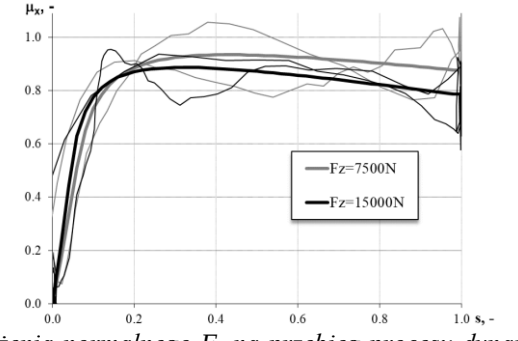
a) przebiegi zmian wartości poślizgu wzdłużnego s_x



b) przebiegi zmian wartości przenoszonej reakcji wzdłużnej F_x



c) porównanie charakterystyk hamowania koła



Rys. 2. Wpływ obciążenia normalnego F_z na przebieg procesu dynamicznego hamowania koła w warunkach laboratoryjnych (koło samochodu ciężarowego średniej ładowności, v_0=60km/h, podłoże typu bęben stalowy gładki)

Są to istotne wnioski z punktu widzenia bezpieczeństwa ruchu pojazdu oraz z punktu widzenia procesu rekonstrukcji zdarzenia drogowego, w którym miało miejsce awaryjne hamowanie pojazdu [15,20,19]. Przedstawiony problem dotyczy głównie samochodów ciężarowych, które mogą przewozić ładunek o masie wyższej niż ich masa własna.

Wydłużenie drogi zatrzymania pod wpływem wzrostu masy pojazdu można oszacować metodą symulacyjną. Zaplanowano i wykonano badania symulacyjne, których celem było przedstawienie wpływu wzrostu masy

pojazdu, a tym samym obciążenia normalnego jego kół jezdnych na wydłużenie drogi zatrzymania w procesie hamowania awaryjnego na podstawie wyników badań eksperymentalnych ogumienia wykonanych w warunkach laboratoryjnych.

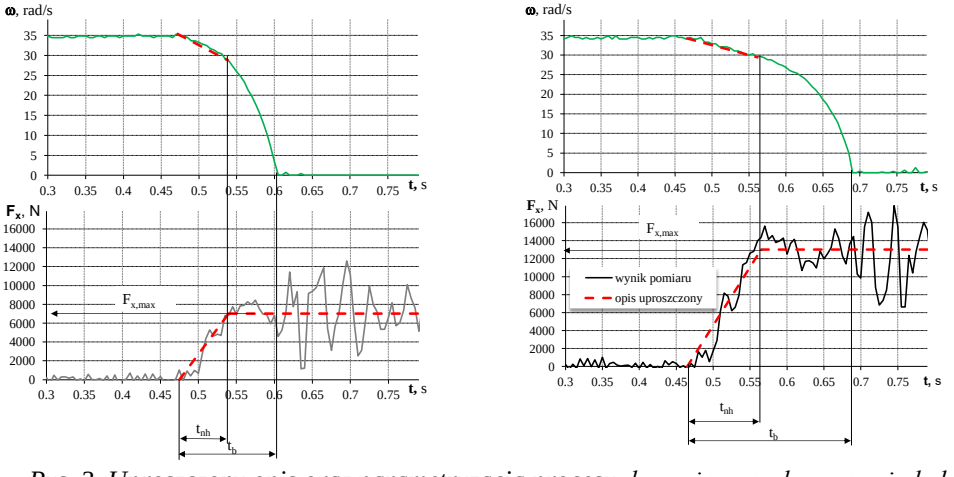
2. Wpływ obciążenia normalnego na przebieg procesu narastania wartości wzdłużnej reakcji stycznej przenoszonej przez koło

Wykonanie zaplanowanych badań symulacyjnych wymagało przygotowania i sparametryzowania uproszczonego opisu procesu narastania wartości wzdłużnej reakcji stycznej przenoszonej przez koło w procesie dynamicznego hamowania.

Proces narastania reakcji wzdłużnej F_x podczas dynamicznego hamowania koła można opisać w sposób uproszczony zależnością liniową, z wykorzystaniem następujących wielkości (Rys. 3):

- wartość graniczna reakcji wzdłużnej $F_{x,max}$ osiągnięta i utrzymana w procesie hamowania koła,
- czas narastania t_{nh} reakcji wzdłużnej F_x.

a) F_z=7500 N b) F_z=15000N

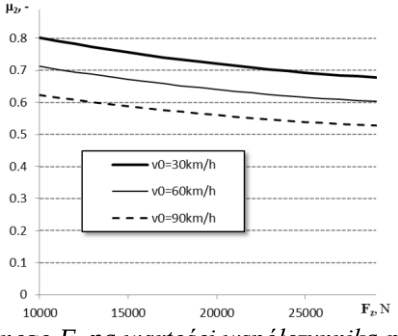


Rys. 3. Uproszczony opis oraz parametryzacja procesu dynamicznego hamowania koła

W czasie narastania t_{nh} reakcji wzdłużnej zachodzi również spadek prędkości kątowej koła ω, a tym samym wzrost wartości poślizgu wzdłużnego s_x (Rys. 3). Dla celów prowadzonej analizy wartość graniczną $F_{x,max}$ reakcji wzdłużnej F_x można ustalić na podstawie współczynnika przyczepności koła do podłoża. Wyniki własnych badań eksperymentalnych koła ogumionego w warunkach

hamowania pokazanych na rysunku 2 oraz przedstawionych w literaturze [7,8] wykazują, że obciążenie normalne koła wpływa na wartość współczynnika przyczepności pojedynczego koła do podłoża. Wraz ze wzrostem obciążenia normalnego koła zmniejsza się zarówno wartość współczynnika przyczepności przylgowej μ_1 jak i poślizgowej μ_2. Można przyjąć, że w procesie hamowania awaryjnego, gdy koła pojazdu ulegają blokowaniu, realna wartość współczynnika przyczepności rozkłada się wokół wartości współczynnika przyczepności poślizgowej μ_2 (Rys. 3).
Na rysunku 3 widać również, że wartość graniczna reakcji wzdłużnej $F_{x,max}$ została osiągnięta w czasie znacznie krótszym niż czas zablokowania koła t_b. Z punktu widzenia prowadzonej analizy, istotnym faktem jest to, że w obu obserwowanych przypadkach styczna reakcja wzdłużna F_x osiągała umowną wartość graniczną $F_{x,max}$, odpowiadającą sile przyczepności poślizgowej, gdy prędkość kątowa koła zmniejszała się od początkowej prędkości kątowej $\omega=\omega_0$ do wartości równiej około $\omega\approx 4/5\omega_0$. Podobne względne spadki prędkości kątowej występują podczas hamowania koła z poziomu innych wartości prędkości początkowej.

Wpływ obciążenia normalnego na przyczepność koła do podłoża w stanie jego zablokowania z uwzględnieniem prędkości toczenia koła przedstawiono na rysunku 4. Wartości współczynnika przyczepności poślizgowej μ_2 koła do podłoża, wyznaczone w szerokim zakresie zmian wartości obciążenia normalnego koła, można bezpośrednio wykorzystać do oszacowania długości drogi zatrzymania pojazdu na podstawie wyrażenia (2).

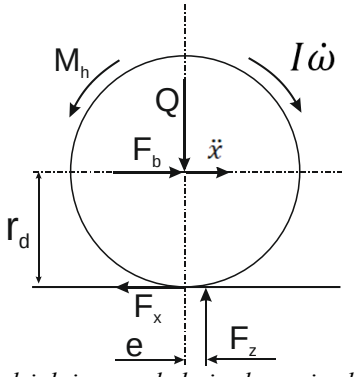


Rys. 4. Wpływ obciążenia normalnego F_z na wartości współczynnika przyczepności poślizgowej μ_2 koła do podłoża (wyniki uzyskane na gładkim podłożu bębna stalowego w warunkach laboratoryjnych)

Na podstawie wniosków z przeprowadzonej obserwacji wyników pomiaru, dla celów realizacji zaplanowanych badań przyjęto następujące założenia upraszczające:

- w procesie hamowania dynamicznego na koło działa maksymalny moment hamowania M_h zadany skokowo, a wartość tego momentu wynika z możliwości mechanizmu sterowania hamulcem koła,
- czas narastania momentu hamowania wynikający z bezwładności mechanizmu sterowania (t_{sh}≈0.2s) nie zależy od obciążenia normalnego koła – czas ten uwzględniono w procesie symulacji, jako stały składnik czasu narastania siły hamowania $t_n = t_{sh} + t_{nh}$,
- w czasie od $t=0$ do t_{nh} ruch obrotowy koła jest jednostajnie opóźniony, zatem prędkość kątowa ω zmienia się liniowo od wartości początkowej $\omega = \omega_0$ (wynikającej z początkowej prędkości toczenia v_0 i promienia dynamicznego koła r_d) do prędkości kątowej $\omega \approx \frac{4}{5}\omega_0$ (Rys. 3),
- w czasie narastania t_{nh}, wzdłużna reakcja styczna przenoszona przez koło narasta liniowo od wartości $F_x=0$ do wartości granicznej $F_{x,max}$ według zależności $F_x(t) = \frac{F_{x,max}}{t_{nh}} t$ (Rys. 3),
- wartość graniczna wzdłużnej rekcji stycznej ograniczona przyczepnością poślizgową koła do podłoża we zależności $F_{x,max} = \mu_2 F_z$ (Rys. 3),
- wartość współczynnika przyczepności poślizgowej koła do podłoża μ_2 jest zależna od obciążenia normalnego koła F_z według zależności pokazanych na rysunku 4.

Przyrost czasu narastania t_{nh} wzdłużnej reakcji stycznej F_x wywołany wzrostem obciążenia normalnego koła można oszacować na podstawie analizy dynamiki ruchu obrotowego koła podczas dynamicznego hamowania. Na rysunku 5 przedstawiono schemat sił i momentów działających na hamowane koło [2].



Rys. 5. Siły i momenty działające na koło jezdne pojazdu podczas hamowania

gdzie:

M_h – moment hamowania koła,

I – sumaryczny moment bezwładności koła wraz z elementami wirującymi związanymi z kołem,

Q – siła obciążająca koło, część ciężaru pojazdu przypadającego na koło ,

F_z – reakcja normalna działająca na koło od nawierzchni drogi,

F_x – reakcja wzdłużna działająca na koło od nawierzchni drogi,

ω – prędkość kątowa koła,

$\ddot{x}$ – przyśpieszenie w ruchu postępowym koła,

F_b– siła popychająca koło, część siły oporu bezwładności pojazdu podczas hamowania.

Na podstawie schematu przedstawionego na rysunku 5 zapisano równanie dynamiki koła w ruchu obrotowym:

$$I\dot{\omega} = -M_h - F_z e + F_x r_d \tag{3}$$

Może się wydawać, że opóźnienie kątowe koła podczas hamowania, a tym samym czas t_{nh} osiągnięcia granicznej wartości reakcji wzdłużnej F_x są w liniowej relacji z obciążeniem normalnym koła F_z. Jednak w warunkach hamowania, na skutek działania reakcji wzdłużnej F_x, opona ulega odkształceniu w kierunku wzdłużnym zgodnie ze zwrotem tej siły. Dlatego równanie powinno być uzupełnione o udział ugięcia wzdłużnego opony. Przebieg charakterystyki sprężystości obwodowej współczesnych opon jest zbliżony do liniowego. Dlatego dla celów obliczeń szacunkowych wartość ugięcia wzdłużnego opony, poczynając od wartości u_x=0, można wyrazić w postaci uproszczonej jako [5,10]:

$$u_x = \frac{F_x}{c_o} \tag{4}$$

c_o – sztywność obwodowa opony

Po uzupełnieniu, równanie dynamiki koła przybiera postać:

$$I\dot{\omega} = -M_h - F_z(e - \frac{F_x}{c_o}) + F_x r_d \tag{5}$$

Po wprowadzeniu dodatkowych założeń uzupełniających:

- w stanie toczenia swobodnego koła przesunięcie reakcji normlanej e jest związane ze współczynnikiem oporu toczenia koła poprzez zależność

$$e = f_t r_d \tag{6}$$

- ugięcie wzdłużne opony u_x narasta liniowo wraz z wartością wzdłużnej reakcji stycznej F_x według zależności

$$u_x(t) = \frac{u_{xmax}}{t_{nh}} t = \frac{F_{x,max}}{c_o t_{nh}} t \tag{7}$$

równanie (5) przybiera postać:

$$I\frac{d\omega}{dt} = -M_h - F_z\left(f_t r_d - \frac{\mu_2 F_z}{c_o t_{nh}} t\right) + \frac{\mu_2 F_z r_d}{t_{nh}} t \tag{8}$$

Poprzez kolejne przekształcenia równania z uwzględnieniem przyjętych wcześniej założeń upraszczających można wyznaczyć czas narastania t_{nh} reakcji wzdłużnej F_x:

$$Id\omega = (-M_h - F_z f_t r_d + F_z \frac{\mu_2 F_z}{c_o t_{nh}} t + \frac{\mu_2 F_z r_d}{t_{nh}} t)dt \tag{9}$$

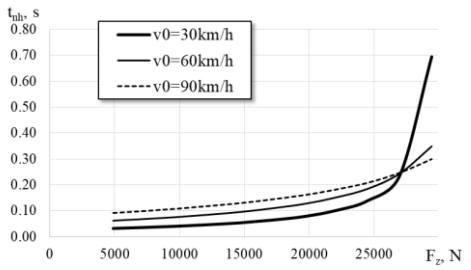
$$\int_{\omega_0}^{\frac{4\omega_0}{5}} Id\omega = \int_0^{t_{nh}} (-M_h - F_z f_t r_d + F_z^2 \frac{\mu_2}{c_o t_{nh}} t + F_z \frac{\mu_2 r_d}{t_{nh}} t)dt \tag{10}$$

$$I\frac{\omega_0}{5} = M_h t_{nh} + F_z f_t r_d t_{nh} - {F_z}^2 \frac{\mu_2}{2c_o} t_{nh} - F_z \frac{\mu_2 r_d}{2} t_{nh} \tag{11}$$

$$I\frac{\omega_0}{5} = t_{nh}\left(M_h + F_z f_t r_d - {F_z}^2 \frac{\mu_2}{2c_o} - F_z \frac{\mu_2 r_d}{2}\right) \tag{12}$$

$$t_{nh} = \frac{I\omega_0}{5\left(M_h + F_z f_t r_d - {F_z}^2 \frac{\mu_2}{2c_o} - F_z \frac{\mu_2 r_d}{2}\right)} \tag{13}$$

Wyniki oszacowania czasu narastania t_{nh} reakcji wzdłużnej F_x do wartości granicznej $F_{x,max}$ przedstawiono na rysunku 6.



Rys 6. Wpływ obciążenia normalnego koła F_z i prędkości toczenia v_0 na wartości czasu narastania t_{nh} siły hamowania koła F_x do wartości granicznej Fx,max (wyniki uzyskane na podstawie symulacji, z wykorzystaniem danych typowych dla samochodu ciężarowego średniej ładowności)

Wartość momentu bezwładności koła z oponą o rozmiarze 275/70R22.5 wraz z piastą i tarczą hamulcową oraz wartość momentu hamowania przyjęto na podstawie wyników badań przedstawionych w innych pracach [7,6].
Wyrażenie (13) pokazuje zależność czasu narastania t_{nh} reakcji wzdłużnej F_x od wielu współczynników charakteryzujących warunki ruchu koła podczas hamowania. Widać bezpośrednią zależność czasu narastania t_{nh} reakcji wzdłużnej F_x od początkowej prędkości toczenia koła v_0 (poprzez związek z prędkością kątową koła ω_0), mimo że wzrost prędkości początkowej toczenia koła v_0 jest częściowo kompensowany spadkiem wartości współczynnika przyczepności koła do podłoża μ_2. W efekcie wyniki przeprowadzonej symulacji wykazały, że czas narastania t_{nh} reakcji wzdłużnej F_x do wartości granicznej $F_{x,max}$ rośnie nieliniowo wraz ze wzrostem obciążenia normalnego koła F_z (Rys. 6). Zależność długości odcinka czasu t_{nh} osiągnięcia przez koło granicznej wartości reakcji wzdłużnej $F_{x,max}$ jest szczególnie nieliniowa w zakresie wysokich wartości obciążenia normalnego koła. Tu widoczny jest udział podatności obwodowej opony, która poprzez sztywność obwodową c_o występującą w równaniu (13) może spowodować znaczne wydłużenie szacowanego czasu t_{nh} szczególnie w zakresie wysokich wartości reakcji wzdłużnej F_x, którym sprzyja występowanie wysokiej wartości obciążenia normalnego koła F_z.

Długości odcinka czasu narastania t_{nh} reakcji wzdłużnej F_x nie są znaczne. Wykazano jednak istotne wydłużenie czasu t_{nh} pod wpływem wzrostu obciążenia normalnego F_z koła. Takie wydłużenie powoduje wydłużenie odcinka czasu narastania siły hamowania t_n pojazdu. Może więc przyczyniać się do zwiększenia długości drogi zatrzymania pojazdu zgodnie z przedstawionym wcześniej równaniem (2).

3. Ocena wpływu masy pojazdu na długość drogi zatrzymania

Obciążenie normalne kół jezdnych pojazdu zmienia się wraz ze zmianą jego masy. Znaczne zmiany masy mogą dotyczyć szczególnie samochodów ciężarowych, gdy dopuszczalna ładowność może nawet dwukrotnie przekraczać masę własną pojazdu.

Wyniki badań eksperymentalnych ogumienia samochodu ciężarowego oraz wyniki obliczeń wykorzystano do symulacji procesu hamowania awaryjnego pojazdu ciężarowego o zmiennej masie. Oprócz wcześniej przedstawionych założeń, przed przygotowaniem aplikacji obliczeniowej przyjęto następujące założenia upraszczające:

- pojazd porusza się na 4 kołach o porównywalnych właściwościach,
- masa pojazdu rozkłada się równomiernie na poszczególne koła,
- kierowca rozpoczyna proces hamowania awaryjnego od chwili zauważenia zagrożenia t=0,
- na każde z kół działa moment hamowania M_h o jednakowej wartości maksymalnej, wynikającej z możliwości mechanizmu sterowania i hamowania,
- intensywność hamowania jest ograniczona współczynnikiem przyczepności poślizgowej μ_2 kół jezdnych do podłoża, o wartości określonej w chwili początku hamowania dla prędkości początkowej v_0 oraz w zależności od obciążenia normalnego każdego z kół (nie zmieniano jego wartości w trakcie hamowania w przyjętym uproszczonym modelu pojazdu),
- współczynnik przyczepności poślizgowej każdego z kół do podłoża jest taki sam, a wynika z ciężaru pojazdu oraz prędkości początkowej hamowania,
- promień dynamiczny każdego z kół jest taki sam,
- nie uwzględniono zmiany obciążenia normalnego osi jezdnych podczas hamowania,
- przyjęto typowe wartości czasu reakcji kierowcy oraz układu hamulcowego [20,4], jednak te jako stałe nie wpływają na obserwowane zjawiska.

Przygotowano arkusz kalkulacyjny umożliwiający wykonanie zaplanowanych obliczeń. Symulację procesu hamowania awaryjnego pojazdu wykonano na bazie równania (2) oraz danych do modelowania opisanych w rozdziale 2. Jednak przebieg zmian prędkości jazdy oraz przebytej drogi wyznaczano w trybie obliczeń iteracyjnych, wyznaczając w określonych

odstępach czasu wartości wielkości fizycznych charakteryzujących przebieg procesu hamowania, w tym:

- opóźnienia hamowania a_h,
- prędkości jazdy v,
- przebytej drogi s.

Zgodnie z założeniami, każdy zestaw wyników był uzyskany dla określonych wartości współczynnika przyczepności poślizgowej μ_2 kół pojazdu do podłoża, z uwzględnieniem masy pojazdu m oraz prędkości początkowej hamowania v_0.
Obliczenia wykonano w następujących wariantach:
- początkowa prędkość jazdy v_0=*30, 60, 90 km/h*,
- masa pojazdu m=4000, 8000, 12000 kg (odpowiednio masa własna pojazdu oraz masa częściowo i całkowicie załadowanego pojazdu ciężarowego do jego DMC),
- kąt nachylenia podłoża $\alpha=0^O$ (podłoże poziome).

Przykładowy zestaw wyników obliczeń przedstawiono na rysunku 7.

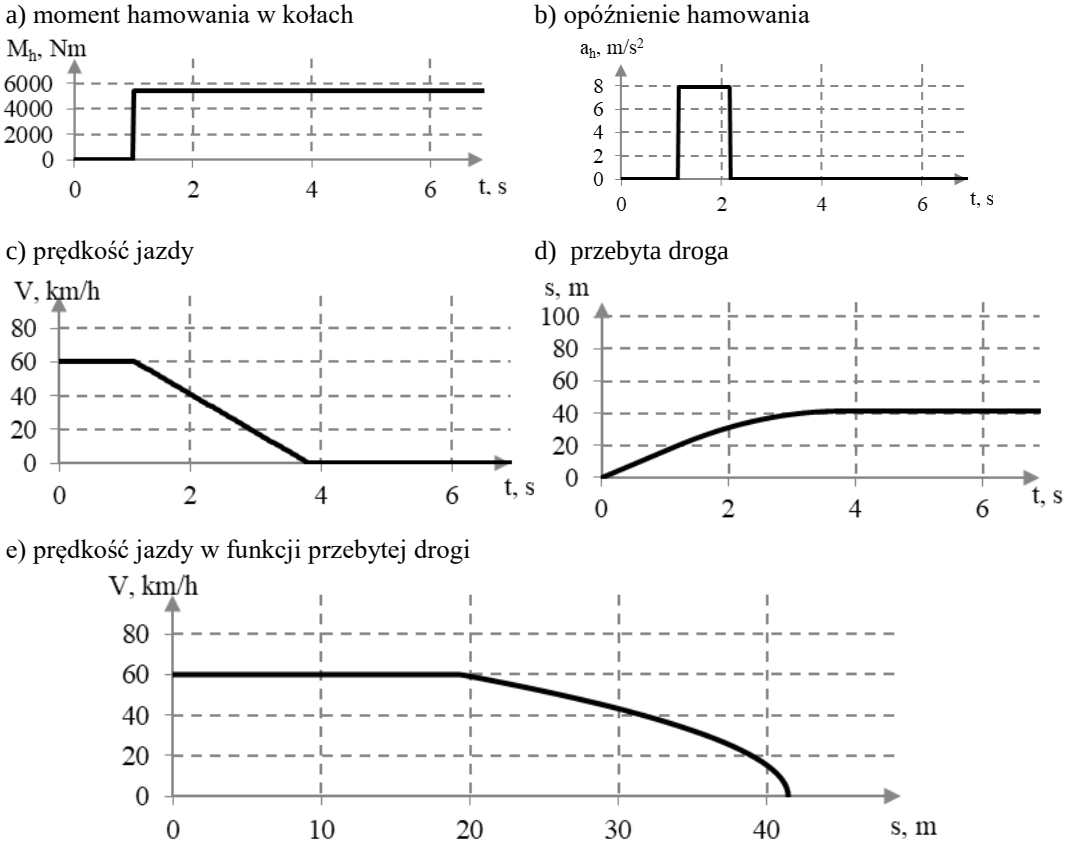
Dla celów prowadzonej analizy największe znaczenie ma dodatkowo wyznaczony przebieg zmian prędkości pojazdu v w funkcji przebytej drogi s, pokazany na rysunku 7.e.

Na przykładzie zestawienia wyników obliczeń pokazanych na rysunku 8 widać wyraźny wpływ zmiany masy pojazdu na wielkości charakteryzujące proces jego hamowania. Zgodnie z oczekiwaniami wyniki obliczeń pokazały, że wzrost masy pojazdu, a więc wzrost obciążenia normalnego każdego z kół wywołał:

- przesunięcie w czasie początku fazy hamowania (Rys. 8.a),
- zmniejszenie intensywności hamowania pojazdu i wydłużenie czasu hamowania (Rys. 8.a, b),
- wydłużenie drogi zatrzymania pojazdu (Rys. 8.c).

Jednak najbardziej wymowny w skutkach wpływ zwiększenia masy pojazdu widać na rysunku 8.d. Przy nieznacznej początkowej prędkości jazdy (v_0=60km/h), wzrost masy pojazdu od masy własnej do masy z częściowym i pełnym obciążeniem ładunkiem spowodował wydłużenie drogi zatrzymania odpowiednio o ok. 4 m i 9 m. Takie wartości wydłużenia drogi zatrzymania są porównywalne odpowiednio z długością typowych pasów przejścia dla pieszych oraz połową długości typowego zestawu siodłowego.
Jednak, z punktu widzenia bezpieczeństwa pojazdu oraz rekonstrukcji zdarzenia drogowego, istotne znaczenie ma zestawienie wyników na rysunku 8.d. Wyniki obliczeń pokazały, że w miejscu gdzie pojazd nieobciążony (m=4000kg)

zatrzymał się, prędkość pojazdu obciążonego w połowie (m=8000kg) wynosiła jeszcze około ponad 20 km/h, a prędkość pojazdu całkowicie obciążonego wynosiła ponad 30 km/h.



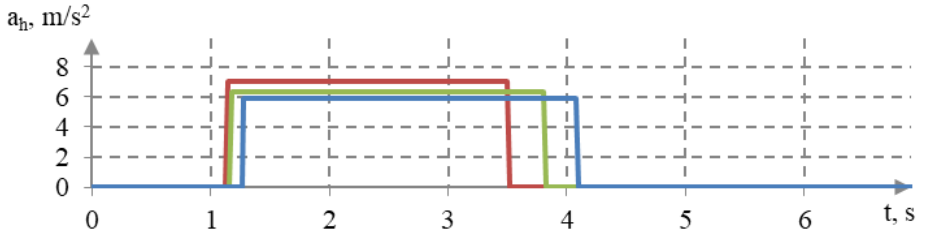
Rys. 7. Przykładowy zestaw wyników obliczeń wielkości fizycznych charakteryzujących proces hamowania awaryjnego pojazdu (v_0=60 km/h, m=8000 kg)

Mimo umiarkowanej początkowej prędkości jazdy, są to znaczne prędkości, których skutki w chwili uderzenia człowieka lub zderzenia pojazdu z dowolnym obiektem mogą być poważne.

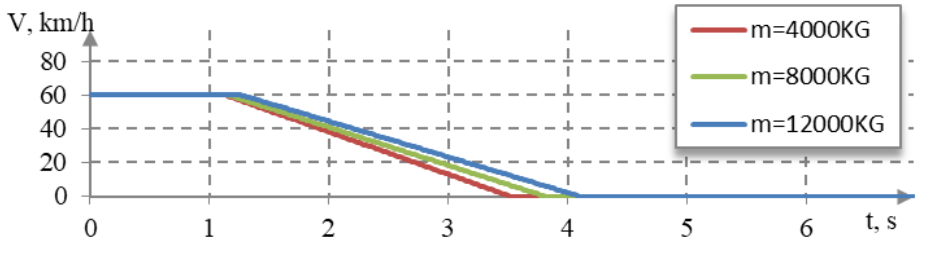
Badania wykazały nieznaczny wpływ zmiany masy pojazdu na czas rozpoczęcia procesu hamowania (Rys. 9). Zgodnie z wynikami obliczeń przedstawionymi na rysunku 6, czas narastania siły hamowania wydłużył się pod wpływem wzrostu masy pojazdu szczególnie w zakresie niskich wartości prędkości początkowej hamowania v_0.

Natomiast wyraźny jest wpływ zmiany masy pojazdu na wartości opóźnienia hamowania, a tym samym na osiągane długości drogi hamowania i zatrzymania.

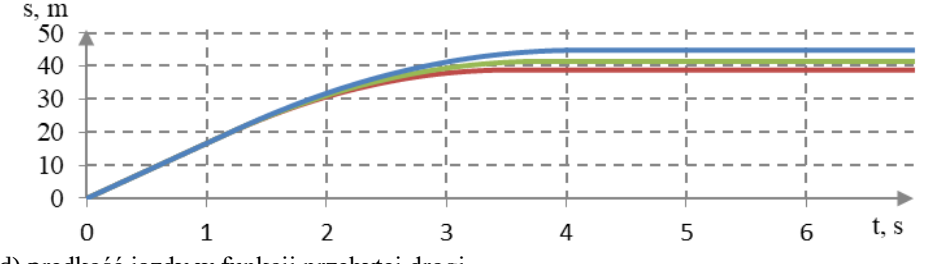
a) opóźnienie hamowania



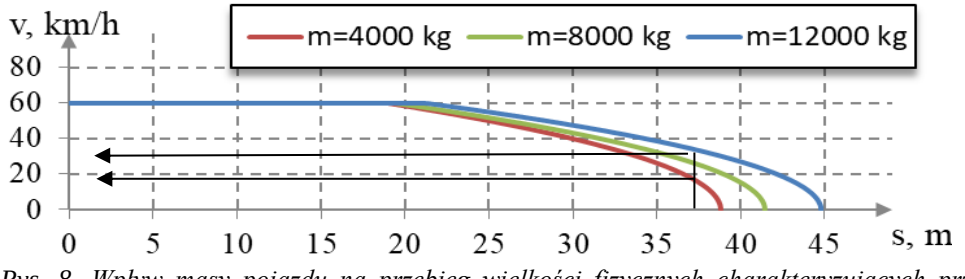
b) prędkość jazdy



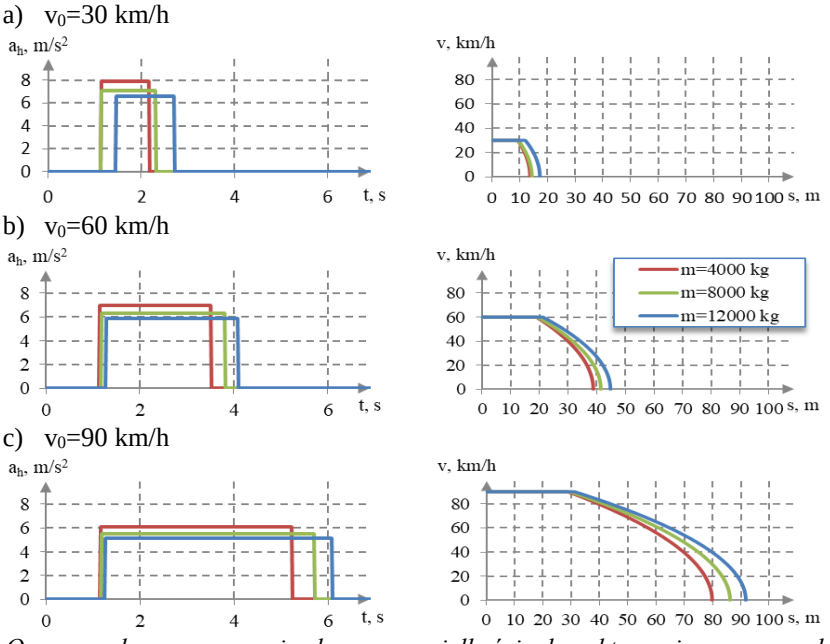
c) przebyta droga



d) prędkość jazdy w funkcji przebytej drogi



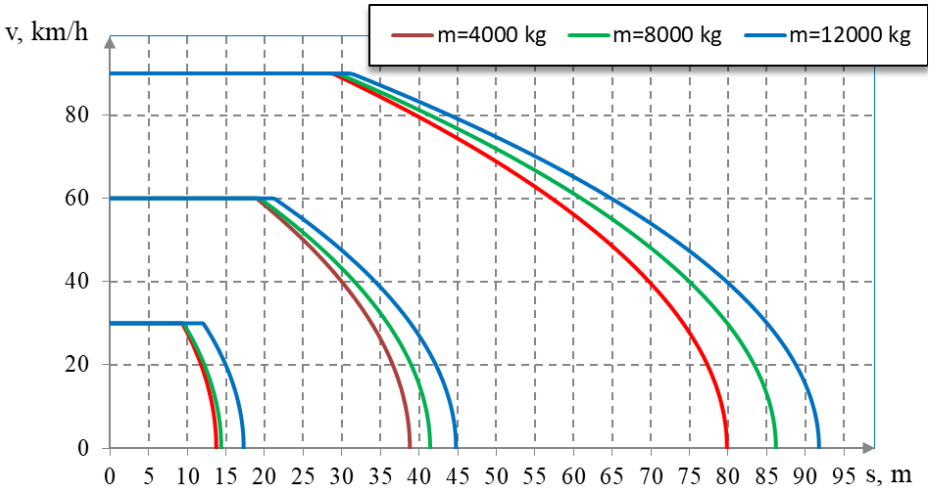
Rys. 8. Wpływ masy pojazdu na przebieg wielkości fizycznych charakteryzujących przebieg procesu hamowania awaryjnego (v_0=60 km/h)



Rys. 9. Ocena wpływu masy pojazdu m na wielkości charakteryzujące proces hamowania awaryjnego przy różnych prędkościach początkowych hamowania

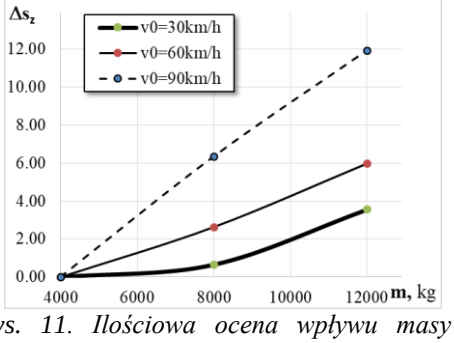
Zestawienie wpływu masy pojazdu na wynik hamowania awaryjnego w warunkach różnej prędkości początkowej hamowania v_0 pokazano na rysunku 10. Przy każdej prędkości początkowej hamowania wzrost masy pojazdu istotnie wydłuża drogę zatrzymania. Jednocześnie widać, że prędkość pojazdu całkowicie obciążonego (m=12000 kg) w miejscu gdzie pojazd nieobciążony (m=4000 kg) ulega zatrzymaniu może wynosić od około *25* km/h do nawet *40* km/h w zależności od prędkości początkowej hamowania.
Na przykładzie uzyskanych wyników badań wykazano, że (Rys. 11):

- przyrost długości drogi zatrzymania pojazdu, na skutek wzrostu jego masy, jest największy w zakresie wysokich prędkości początkowej hamowania v_0,
- w ujęciu względnym, przyrost masy pojazdu od masy własnej do DMC może spowodować wydłużenie drogi zatrzymania pojazdu nawet o ponad 20%.

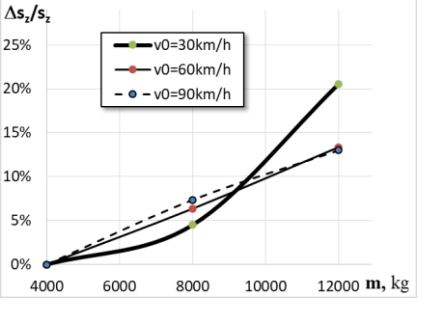


Rys. 10. Ocena wpływu masy pojazdu m długość drogi zatrzymania przy różnych prędkościach początkowych hamowania v_0

a) przyrost bezwzględny



b) przyrost względny



Rys. 11. Ilościowa ocena wpływu masy pojazdu m na przyrost drogi zatrzymania s_z w warunkach hamowania awaryjnego

4. Wnioski końcowe

Wykonane badania symulacyjne wykazały, że zwiększenie masy pojazdu może istotnie wydłużyć drogę zatrzymania awaryjnego pojazdu poprzez:

- zwłokę początku procesu hamowania,
- zmniejszenie intensywności hamowania.

Ponadto wykazano, że pojazd obciążony ładunkiem, może poruszać się ze znaczną prędkością w chwili gdy pojazd nieobciążony ulega zatrzymaniu. Są to istotne wnioski z punktu widzenia bezpieczeństwa ruchu pojazdu. Jednocześnie

wnioski te wskazują na konieczność uwzględniania tak istotnych zmian własności ogumienia w procesie analizy i rekonstrukcji zdarzenia drogowego.

Możliwa jest kontynuacja badań zarówno w części eksperymentalnej jak i modelowej. Pokazano, że u podstaw obserwowanych zjawisk zmian leżą ujawnione podczas badań eksperymentalnych własności koła ogumionego. Jednak w badaniach symulacyjnych przyjęto pewne założenia upraszczające, w tym uproszczony model tarcia zablokowanego koła względem podłoża, z ustaloną wartością współczynnika przyczepności w cały procesie hamowania. Spodziewany jest większy wpływ wzrostu masy pojazdu na wydłużenie jego drogi zatrzymania, gdy w badaniach zostaną uwzględnione:

- realne zmiany wartości współczynnika przyczepności poślizgowej koła do podłoża pod wpływem prędkości,
- obniżenie wartości współczynnika przyczepności koła do podłoża w znacznie wydłużonym czasie jego hamowania.

Te zagadnienia mogą stanowić główne kierunki dalszych badań.

Literatura

1. Andrzejewski R. Dynamika pneumatycznego koła jezdnego. WNT, 2010 Warszawa.
2. Arczynski S. Mechanika ruchu samochodu. Warszawa: WPW, 1984r.
3. Goudie D, Bowler J, Brown C, Heinrichs B. et al. Tire Friction During Locked Wheel Braking. 2000. SAE Technical Paper 2000-01-1314.
4. Jurecki R, Jaskiewicz M, Guzek M, Lozia Z, Zdanowicz P. Driver's reaction time under emergency braking a car – Research in a driving simulatory. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (4): 295–301.
5. Kulikowski K, Szpica D. Determination of directional stiffnesses of vehicles' tires under a static load operation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (1): 66–72.
6. Lawniczak S, Prochowski L. Analiza zmian predkosci katowej kola samochodu ciezarowego podczas hamowania Zeszyty Instytutu Pojazdow/Politechnika Warszawska 2001; z. 1/40: 83-92.
7. Luty W, Prochowski L. Analiza procesu narastania sily hamowania kola samochodu ciezarowego. Zeszyty Instytutu Pojazdow/Politechnika Warszawska 2001; 1(40): 9-24.

8. Luty W, Prochowski L. Modelowanie charakterystyk przyczepnosci ogumienia samochodow ciezarowych. Zeszyty Instytutu Pojazdow/ Politechnika Warszawska 2002; 1(44): 37-47.
9. Luty W. Analiza własciwosci nowych konstrukcji ogumienia kołjezdnych samochodu ciezarowego sredniej ladownosci. Praca zbiorowa pt. Analiza wpływu ogumienia nowych konstrukcji na bezpieczenstwo samochodu w ruchu krzywoliniowym. Warszawa: Wojskowa Akademia Techniczna, 2009; 7-18.
10. Luty W.: Badania eksperymentalne oraz opis analityczny własciwosci ogumienia samochodow. Zeszyty Naukowe Instytutu Pojazdow Politechniki Warszawskiej 2010; 1(77): 7-26.
11. Milliken W F, Milliken D L. Race Car Vehicle Dynamics, Warrendale 1995; SAE International.
12. Mitschke M. Teoria samochodu. Dynamika samochodu (Tom 1: Naped i hamowanie). Warszawa 1987; WKŁ.
13. Parczewski K, Wnek H. Make use of the friction coefficient during braking the vehicle. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14 (2): 176-180.
14. Parczewski K. Effect of tyre inflation pressure on the vehicle dynamics during braking manouvre. Eksploatacja i Niezawodnosc. Maintenance and Reliability 2013; 15 (2): 134–139.
15. Prochowski L, Unarski J, Wach W, Wicher J. Podstawy rekonstrukcji wypadkow drogowych. Warszawa 2008; WKŁ.
16. Prochowski L. Pojazdy samochodowe. Mechanika ruchu. Warszawa 2005; WKŁ.
17. Reed W. Keskin A. Vehicular Deceleration and Its Relationship to Friction. 1989. SAE Technical Paper 890736.
18. Sharizli A, et al. Simulation and Analysis on the Effect of Gross Vehicle Weight on Braking Distance of Heavy Vehicle, Applied Mechanics and Materials 2014, Vol. 564, pp. 77-82.
19. Warner C, Smith G, James M, Germane G. Friction Applications in Accident Reconstruction. 1983. SAE Technical Paper 830612.

20. Wiercinski J, Reza A, i inni. Wypadki drogowe. Krakow 2010; Vademecum biegłego sadowego. Wydanie 2 uaktualnione, IES.

Dr hab. inż. Mariusz ZIEJA

Instytut Techniczny Wojsk Lotniczych
ul. Księcia Bolesława 6, 01-494 Warszawa 96
e-mail: mariusz.zieja@itwl.pl

Dr inż. Mariusz WAŻNY

Wydział Mechatroniki i Lotnictwa
Wojskowa Akademia Techniczna
ul. Kaliskiego 2, 00-908 Warszawa 49
e-mail: mwazny@wat.edu.pl

Dr inż. Sławomir STĘPIEŃ

Wydział Mechatroniki i Lotnictwa
Wojskowa Akademia Techniczna
ul. Kaliskiego 2, 00-908 Warszawa 49
e-mail: sstepien@wat.edu.pl

Zarys metody szacowania trwałości elementów lub zespołów urządzeń z zachowaniem wymaganego poziomu niezawodności

Słowa kluczowe: *niezawodność, trwałość, funkcja gęstości stan dopuszczalny, starzenie, zużywanie*

Streszczenie: Praca zawiera probabilistyczną metodę oceny trwałości elementów lub zespołów urządzeń pracujących w warunkach oddziaływania procesów destrukcyjnych. W wyniku działania tychże procesów następuje zużywanie powodujące pogorszenie warunków ich współpracy. Przyjmuje się, że element pracuje niezawodnie, gdy zużycie nie przekracza wartości dopuszczalnych (granicznych). Metoda od strony matematycznej bazuje na równaniu różnicowym z którego po przekształceniu otrzymuje się równanie różniczkowe cząstkowe typu Fokkera-Plancka. Z rozwiązania szczególnego tego równania otrzymuje się funkcję gęstości prawdopodobieństwa zużywania w postaci rozkładu normalnego. W pracy przedstawione są dwa sposoby wyznaczania trwałości. Pierwszy polega na wykorzystaniu funkcji gęstości zużywania a drugi na wyznaczeniu funkcji gęstości prawdopodobieństwa czasu osiągania stanu dopuszczalnego i zastosowanie jej do wyznaczenia trwałości elementu lub zespołu. W pracy przedstawiono przykład liczbowy dotyczący procesu eksploatacji techniki lotniczej.

1. Wstęp

W dostępnej literaturze można znaleźć szereg prac, które pokazują problem wpływu środowiska zewnętrznego, procesów starzenia i zużywania na funkcjonowanie systemu technicznego [4,9,13,16,17,21]. Ze względu z na zaawansowanie techniczne i duży stopień integracji urządzeń wykorzystywanych na pokładzie wojskowych statków powietrznych opracowanie optymalnych modeli eksploatacji jest zadaniem skomplikowanym. Niezwykle użyteczne w tym obszarze są metody oceny niezawodności i trwałości urządzeń lotniczych bazujące na zmianie parametrów diagnostycznych [6,7,8,12,15,20].

Niniejsza praca zawiera probabilistyczną metodę oceny trwałości elementów i zespołów urządzenia pracującego w warunkach oddziaływania procesów starzeniowych (korozyjnych, zużyciowych i innych) w urządzeniach statku powietrznego [15,18.19]. Stan techniczny niektórych urządzeń lotniczych można oceniać przy pomocy parametrów diagnostycznych. Ocena ta wymaga znajomości granicznych (dopuszczalnych) wartości dla których uznaje się, że urządzenie lub zespół jest w stanie zdatności do użycia.

W proponowanym modelu oceny trwałości przyjmuje się następujące założenia:

– stan techniczny urządzenia określa się jednym parametrem diagnostycznym „z" w postaci odchyłki parametru od wartości nominalnej,

$$z = \left| X - X^{norm} \right|, \tag{1}$$

gdzie:

X – wartość bieżąca parametru diagnostycznego,

X^{norm} – wartość nominalna parametru diagnostycznego;

– zmiana wartości odchyłki parametru diagnostycznego następuje w całym okresie eksploatacji (pracy i postoju);
– parametr „z" jest niemalejący, ponieważ określony jest przez wartość bezwzględną różnicy wartości bieżącej i wartości nominalnej;
– prędkość narastania odchyłki parametru diagnostycznego można w przypadku zmian losowych opisać następującą zależnością:

$$\frac{dz}{dt} = c\,, \tag{2}$$

gdzie:

c – zmienna losowa charakteryzująca podatność elementu na zmiany starzeniowe w zależności od jego cech i warunków pracy,

t – czas kalendarzowy.

2. Metoda szacowania trwałości elementu urządzenia z wykorzystaniem funkcji gęstości odchyłki parametru diagnostycznego

2.1. Wyznaczenie funkcji gęstości odchyłki z uwzględnieniem zależności (1)

Dynamika zmian wartości odchyłki „z" w ujęciu losowym scharakteryzujemy następującym równaniem różnicowym

$$U_{z,t+\Delta t} = (1-P)U_{z,t} + PU_{z-\Delta z,t}\,, \tag{3}$$

gdzie:

$U_{z,t}$ - prawdopodobieństwo tego, że w chwili t wartość parametru diagnostycznego przyjmuje wartość z;

P - prawdopodobieństwo zdarzenia, że występuje losowe zużywanie i w przedziale czasu o długości Δt wartość odchyłki wzrośnie o wartość Δz;

Δz - przyrost odchyłki.

Dla przypadku, gdy $P=1$ równanie (3) w zapisie funkcyjnym przyjmie postać:

$$u(z, t+\Delta t) = u(z-\Delta z, t)\,. \tag{4}$$

Równanie (4) ma następujący sens: prawdopodobieństwo tego, że w chwili t +Δt wartość odchyłki będzie równa z jest równe prawdopodobieństwu tego, że w chwili t wartość odchyłki była równa z-Δz. Oznacza to, że z prawdopodobieństwem równym jedności, w przedziale czasu o długości Δt, odchyłka wzrośnie o wartość Δz.

Równanie (4) przekształcamy w równanie różniczkowe cząstkowe. W tym celu przyjmujemy następujące przybliżenia [1,2]:

$$u(z,t+\Delta t)=u(z,t)+\frac{\partial u(z,t)}{\partial t}\Delta t\,, \tag{5}$$

$$u(z-\Delta z,t)=u(z,t)-\frac{\partial u(z,t)}{\partial z}\Delta z+\frac{1}{2}\frac{\partial^2 u(z,t)}{\partial z^2}(\Delta z)^2\,. \tag{6}$$

Wykorzystując (5) i (6) to równanie (4) przyjmuje następującą postać:

$$\frac{\partial u(z,t)}{\partial t}=-b\frac{\partial u(z,t)}{\partial z}+\frac{1}{2}a\frac{\partial^2 u(z,t)}{\partial z^2}\,, \tag{7}$$

gdzie:

$b=E[c]$ - średni przyrost wartości odchyłki parametru diagnostycznego na jednostkę czasu;

$a=E[c^2]$ - średni kwadrat przyrostu odchyłki parametru diagnostycznego na jednostkę czasu.

Szukamy rozwiązania szczególnego równania (7) takiego, które przy $t\to 0$ jest zbieżna do tzw. funkcji Diraca, tj. $u(z,t)\to 0$ dla z≠0 i $u(0,t)\to+\infty$, ale w ten sposób, że całka funkcji *u* jest równa „1" dla wszystkich $t>0$.

Rozwiązanie równania (7) dla wyżej określonego warunku przyjmuje postać [3,11,14]:

$$u(z,t)=\frac{1}{\sqrt{2\pi A(t)}}e^{-\frac{(z-B(t))^2}{2A(t)}}\,, \tag{8}$$

gdzie: $B(t)=\int_0^t b\,dt=bt=\bar{c}t,\qquad A(t)=\int_0^t a\,dt=at=\bar{c}^2t\,.$

Wartość 0 w dolnych granicach całek oznacza przyjętą, początkową chwilę czasu, od której rozpatrujemy dynamikę zmian wartości parametru diagnostycznego - może to być np. moment wprowadzenia danego urządzenia do eksploatacji.

Funkcję gęstości (8) narastania odchyłki parametru diagnostycznego można wykorzystać do oceny niezawodności elementu rozpatrywanego urządzenia.

2.2. Określenie niezawodności i trwałości elementu lub zespołu urządzenia

Dysponując określoną funkcją gęstości można zapisać zależność na niezawodność i trwałość ze względu na czas narastania odchyłki parametru do wartości granicznej. Wzór przyjmuje postać

$$R(t)=\int_{-\infty}^{z_d}u(z,t)dz\,, \tag{9}$$

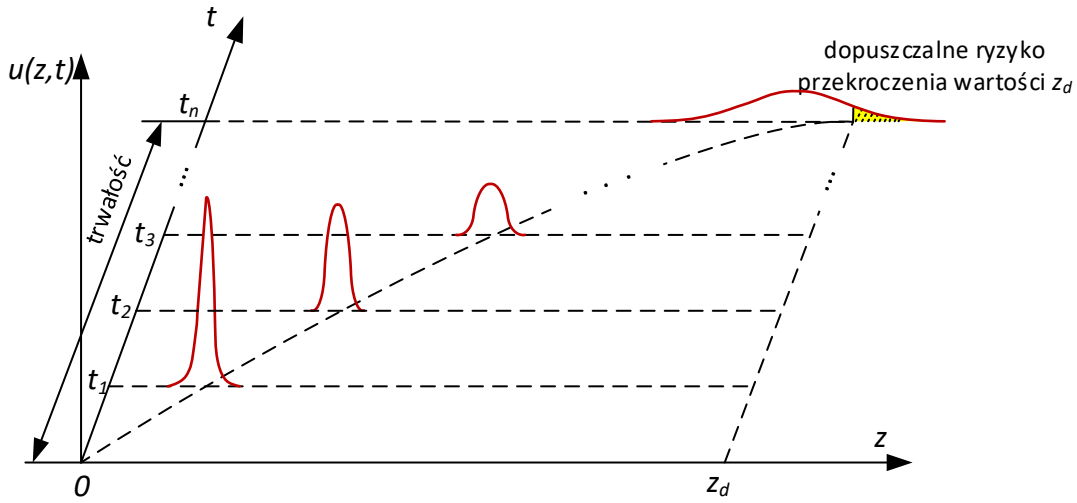
gdzie:

$u(z,t)$ - funkcja gęstości określona zależnością (8);

z_d - dopuszczalna wartość odchyłki parametru diagnostycznego ze względu na bezpieczeństwo;

t - czas kalendarzowy eksploatacji urządzenia.

Na rysunku 1 przedstawiono schemat przebiegu funkcji gęstości i sposobu określenia niezawodności oraz trwałości.



Rys. 1. Schemat zmian postaci funkcji gęstości

Zależność (9) z uwzględnieniem (8), przyjmuje postać:

$$R(t)=\int_{-\infty}^{z_d}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z-bt)^2}{2at}}dz \tag{10}$$

Zakładając minimalną, wymaganą wartość niezawodności R^* można określić czas t^*, po którym niezawodność spadnie poniżej wymaganego poziomu. Czas t^* można traktować jako trwałość danego elementu dla żądanej, dopuszczalnej wartości niezawodności.

W tym przypadku otrzymujemy:

$$R^*=\int_{-\infty}^{z_d}\frac{1}{\sqrt{2\pi at^*}}e^{-\frac{(z-bt^*)^2}{2at^*}}dz \tag{11}$$

3. Metoda szacowania trwałości z wykorzystaniem funkcji gęstości czasu przekroczenia stanu dopuszczalnego (granicznego)

3.1. Wyznaczanie rozkładu czasu przekraczania stanu dopuszczalnego (granicznego)

Prawdopodobieństwo przekroczenia wartości dopuszczalnej (granicznej) przez parametr diagnostyczny z wykorzystaniem funkcji gęstości zmian odchyłki parametru diagnostycznego (8) można przedstawić w postaci:

$$Q(t;z_d)=\int_{z_d}^{\infty}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z-bt)^2}{2at}}dz\,, \tag{12}$$

Funkcję gęstości rozkładu czasu pierwszego przejścia poza wartość dopuszczalną z_d przyjmuje postać:

$$f(t)=\frac{\partial}{\partial t}Q(t;z_d)=\frac{\partial}{\partial t}\int_{z_d}^{\infty}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z-bt)^2}{2at}}dz\,. \tag{13}$$

Zatem,

$$f(t)=\int_{z_d}^{\infty}\left\{\frac{\partial}{\partial t}\left[\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z-bt)^2}{2at}}\right]\right\}dz\,. \tag{14}$$

Przyjmując określenie (8) otrzymujemy:

$$f(t)=\int_{z_d}^{\infty}\left\{\frac{\partial}{\partial t}u(z,t)\right\}dz\,. \tag{15}$$

Ponadto pochodna po czasie funkcji (8) przyjmuje następującą postać

$$\frac{\partial}{\partial t}\left[u(z,t)\right]=u(z,t)\left(\frac{z^2-b^2t^2-at}{2at^2}\right). \tag{16}$$

Zależność (16) podstawiono do (14)

$$f(t)=\int_{z_d}^{\infty}\left[u(z,t)\left(\frac{z^2-b^2t^2-at}{2at^2}\right)\right]dz\,. \tag{17}$$

Szukamy teraz funkcji pierwotnej dla funkcji podcałkowej w zależności (17). Przewidujemy, że funkcja postaci:

$$w(z,t)=u(z,t)\theta(z,t),$$

jest funkcją pierwotną dla funkcji podcałkowej zależności (17), gdzie $\theta(z,t)$ jest poszukiwaną, nieznaną funkcją.

Czyli

$$\frac{\partial}{\partial z}\left[u(z,t)\theta(z,t)\right]=u(z,t)\left(\frac{z^2-b^2t^2-at}{2at^2}\right),$$

Po przekształceniach uzyskujemy następujące równanie:

$$\frac{\partial\theta(z,t)}{\partial z}-\frac{z-bt}{at}\theta(z,t)=\frac{z^2-b^2t^2-at}{2at^2}. \tag{18}$$

Równanie jednorodne:

$$\frac{\partial\theta(z,t)}{\partial z}-\frac{z-bt}{at}\theta(z,t)=0\,.$$

Rozwiązanie równania jednorodnego:

$$\theta_o(z,t)=Ce^{\frac{z^2-2btz}{2at}},$$

gdzie: C - stała dowolna

Przewidywane rozwiązanie szczególne równania niejednorodnego ma postać:

$$\theta_s(z,t)=-\frac{z+bt}{2t}$$

Sprawdzono, że powyższe rozwiązanie spełnia równanie (18). Rozwiązanie ogólne równania niejednorodnego:

$$\theta(z,t)=Ce^{\frac{z^2-2btz}{2at}}-\frac{z+bt}{2t},$$

Czyli poszukiwana funkcja pierwotna całki (17) ma postać:

$$w(z,t)=u(z,t)\left[Ce^{\frac{z^2-2btz}{2at}}-\frac{z+bt}{2t}\right],$$

Zatem obliczając całkę (17) w podanych granicach otrzymujemy

$$f(t)=u(z,t)\left[Ce^{\frac{z^2-2btz}{2at}}-\frac{z+bt}{2t}\right]_{z_d}^{\infty}=Cu(z,t)e^{\frac{z^2-2btz}{2at}}\Big|_{z_d}^{\infty}-u(z,t)\frac{z+bt}{2t}\Big|_{z_d}^{\infty}=$$

$$=C\frac{1}{\sqrt{2\pi at}}e^{-\frac{b^2t}{2a}}\Big|_{z_d}^{\infty}-u(z,t)\frac{z+bt}{2t}\Big|_{z_d}^{\infty}=0-0+u(z_d,t)\frac{z_d+bt}{2t}\Big|$$

$$f(t)=\frac{z_d+bt}{2t}u(z_d,t). \tag{19}$$

Zależność (19) określa funkcję gęstości czasu pierwszego przejścia stanu dopuszczalnego (granicznego) przez odchyłkę parametru diagnostycznego. Należy sprawdzić, czy funkcja (19) jest funkcją gęstości czasu osiągnięcia stanu dopuszczalnego (granicznego). Funkcja ta ma postać:

$$f(t)=\frac{z_d+bt}{2t}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z_d-bt)^2}{2at}}. \tag{20}$$

Funkcja (20) powinna spełniać warunek

$$\int_0^\infty f(t)dt=1. \tag{21}$$

Dla wykazania słuszności (21) przedstawia się następujące uzasadnienia

$$\int_0^\infty \frac{z_d+bt}{2t}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z_d-bt)^2}{2at}}dt=1. \tag{22}$$

W celu obliczenia całki występującej we wzorze (22) stosujemy następujące podstawienie:

$$w=\frac{z_d-bt}{\sqrt{at}}\Rightarrow dt=-\frac{2t\sqrt{at}}{z_d+bt}dw. \tag{23}$$

Przekształcenie granic całkowania

$$t=0\Rightarrow w=\infty$$

$$t=\infty\Rightarrow w=\lim_{t\to\infty}\frac{z_d-bt}{\sqrt{at}}\overset{H}{=}\lim_{t\to\infty}\frac{-2b\sqrt{at}}{a}=-\infty. \tag{24}$$

Po podstawieniu do wyjściowej całki otrzymujemy:

$$\int_0^\infty \frac{z_d+bt}{2t}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z_d-bt)^2}{2at}}dt=-\int_\infty^{-\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw. \tag{25}$$

Powyższa całka jest całką z rozkładu normalnego N(0,1) w granicach od $-\infty$ do $+\infty$ i jest równa jedności. Na tej podstawie można stwierdzić, że:

$$\int_0^\infty \frac{z_d+bt}{2t}\frac{1}{\sqrt{2\pi at}}e^{-\frac{(z_d-bt)^2}{2at}}dt=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2}}dw=1. \tag{26}$$

3.2. Ocena trwałości wybranych elementów konstrukcji statku powietrznego z wykorzystaniem rozkładu czasu osiągania stanu dopuszczalnego

Wzór na niezawodność elementu konstrukcyjnego statku powietrznego przyjmuje postać:

$$R(t)=1-\int_0^t f(\tau)d\tau, \tag{27}$$

gdzie funkcja gęstości $f(t)$ określona jest wzorem (19).

Natomiast, zawodność elementu konstrukcyjnego statku powietrznego można wyznaczyć z zależności

$$Q(t)=\int_0^t \frac{z_d+b\tau}{2\tau}\cdot\frac{1}{\sqrt{2\pi a\tau}}e^{-\frac{(z_d-b\tau)^2}{2a\tau}}d\tau. \tag{28}$$

Całkę występującą w zależności (27) oraz (28) należy przekształcić do dogodniejszej postaci:

$$\int_0^t \frac{z_d + b\tau}{2\tau} \cdot \frac{1}{\sqrt{2\pi a\tau}} e^{-\frac{(z_d - b\tau)^2}{2a\tau}} d\tau = \left| \begin{array}{ll} w = \dfrac{z_d - b\tau}{\sqrt{a\tau}} & \tau = 0 \Rightarrow w = \infty \\ d\tau = -\dfrac{2\tau\sqrt{a\tau}}{z_d + b\tau} dw & \tau = t \Rightarrow w = \dfrac{z_d - bt}{\sqrt{at}} \end{array} \right| = -\int_{\infty}^{\frac{z_d - bt}{\sqrt{at}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw$$

Po zamianie granic całowania otrzymujemy

$$\int_0^t \frac{z_d + b\tau}{2\tau} \cdot \frac{1}{\sqrt{2\pi a\tau}} e^{-\frac{(z_d - b\tau)^2}{2a\tau}} d\tau = \int_{\frac{z_d - bt}{\sqrt{at}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw \tag{29}$$

Niezawodność danego elementu przyjmie postać:

$$R(t) = 1 - \int_{\frac{z_d - bt}{\sqrt{at}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw, \tag{30}$$

lub

$$R(t) = \int_{-\infty}^{\frac{z_d - bt}{\sqrt{at}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw. \tag{31}$$

Całka występująca we wzorze (31) jest wartością dystrybuanty rozkładu normalnego N(0,1) dla argumentu występującego w górnej granicy całkowania. Ponownie zakładając wymaganą, minimalną wartość niezawodności R^* możemy wyznaczyć trwałość t^*.

$$R^* = \int_{-\infty}^{\frac{z_d - bt^*}{\sqrt{at^*}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw. \tag{32}$$

Skorzystanie w obliczeniach ze wzoru (11) lub (32) wymaga oszacowania wartości współczynników *a* oraz *b*. Oszacowania tego dokonuje się na podstawie danych uzyskanych z procesu eksploatacji statku powietrznego.

4. Przykład liczbowy

W celu określenia trwałości rozpatrywanego elementu należy wyznaczyć (oszacować) wartości stałych *a* oraz *b*. W tym celu przyjmujemy, że z obserwacji badanego urządzenia w procesie eksploatacji posiadamy dane o narastaniu wartości odchyłki parametru diagnostycznego w postaci

$$[(z_0, t_0), (z_1, t_1), (z_2, t_2), ..., (z_n, t_n)]. \tag{33}$$

Najlepszą metodą do wyznaczania wartości „*b*" i „*a*" dla posiadanych danych jest metoda wykorzystująca funkcję wiarogodności. Jej postać w ogólnym przypadku możemy przedstawić jako zależność:

$$L = \prod_{k=0}^{n-1} g(t_k, z_k, \theta_1, \theta_2, ..., \theta_m), \tag{34}$$

gdzie:

$g(t_k, z_k, \theta_1, \theta_2, ..., \theta_m)$ – funkcja gęstości prawdopodobieństwa całkowitego zmiennej *z*;

$(\theta_1, \theta_2, ..., \theta_m)$ – parametry funkcji gęstości;

z_k – pomierzone wartości zużycia parametru *z* odpowiednio w chwilach czasu $(t_1, t_2, ..., t_k)$.

Znalezienie oszacowań $\theta_1^*, \theta_2^*, ..., \theta_m^*$ nieznanych parametrów $\theta_1, \theta_2, ..., \theta_m$ metodą największej wiarogodności sprowadza się do rozwiązania równań postaci:

$$\frac{\partial \ln L}{\partial \theta_j} = 0, \tag{35}$$

gdzie:

j=1,2,...,m;

m - liczba parametrów charakteryzujących proces zużycia danego obiektu technicznego.

W tym przypadku oszacowanie b^* i a^* nieznanych parametrów *b* i *a* metodą największej wiarogodności sprowadza się do rozwiązania układu równań:

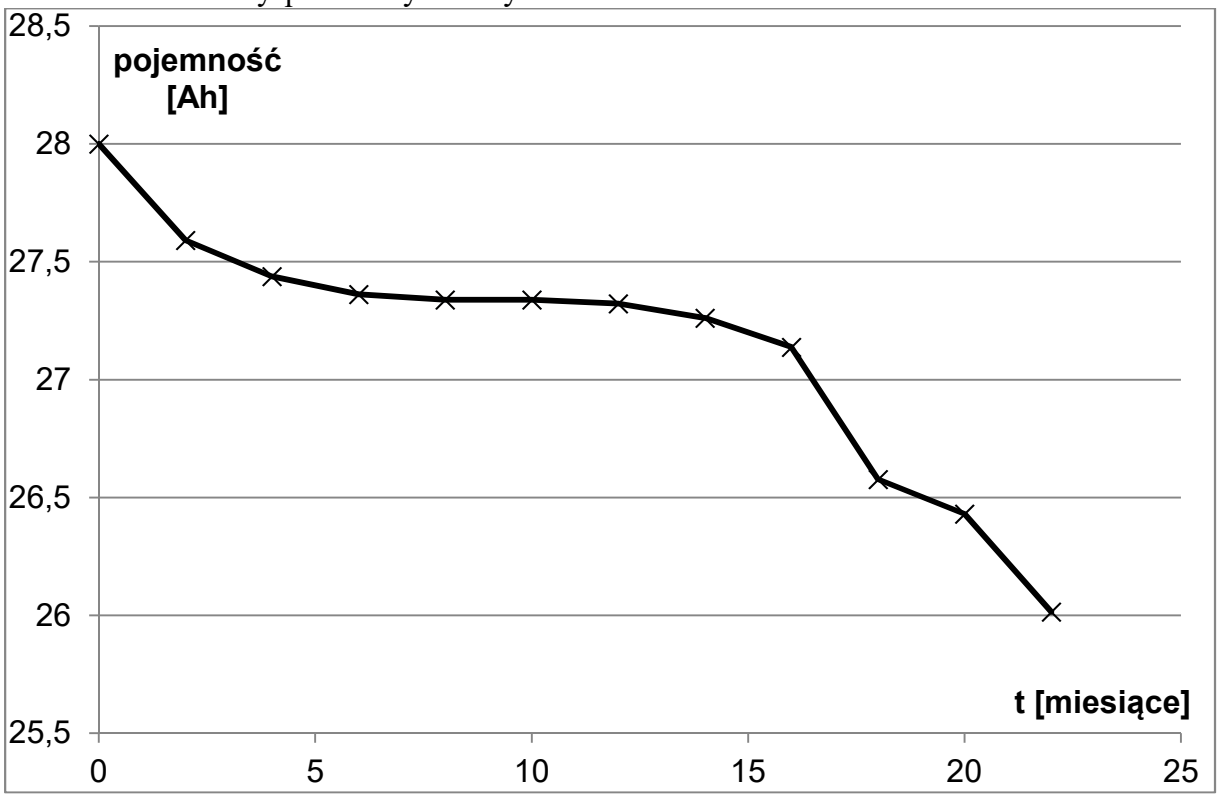
$$\begin{cases} \dfrac{\partial \ln L}{\partial b} = 0 \\ \dfrac{\partial \ln L}{\partial a} = 0 \end{cases}. \tag{29}$$

Rozwiązując układ równań (29) znajdujemy b^* i a^*.

$$b^* = \frac{z_n}{t_n}, \tag{36}$$

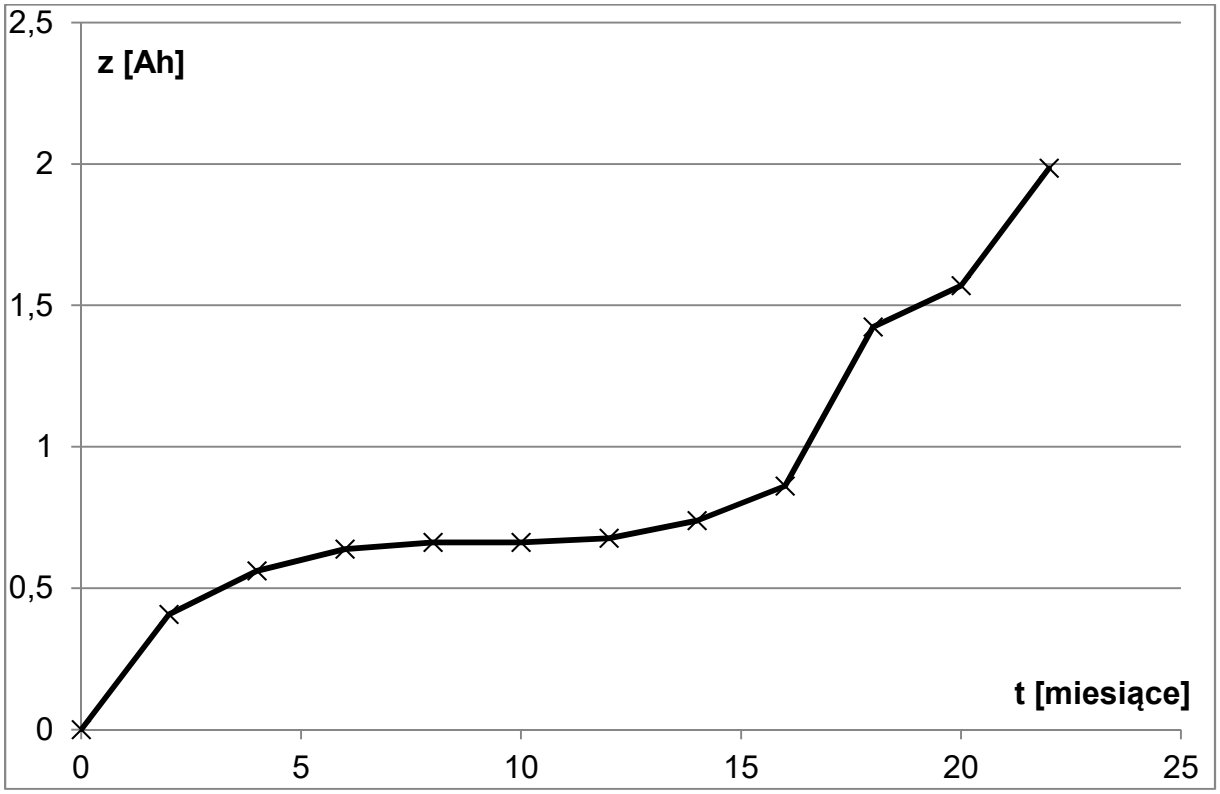
$$a^* = \frac{1}{n}\sum_{k=0}^{n-1} \frac{\left[\left(z_{k+1} - z_k\right) - b^*\left(t_{k+1} - t_k\right)\right]^2}{\left(t_{k+1} - t_k\right)}. \tag{37}$$

Elementem, który został wybrany do przykładu liczbowego jest bateria lotnicza 12-SAM-28. Na rysunku 2 przedstawiono zmianę w czasie uśrednionej pojemności akumulatorów dla dysponowanych danych.



Rys. 2. Przebieg zmian uśrednionej pojemności baterii 12-SAM-28

Za parametr diagnostyczny „*z*” przyjęto, zgodnie z zależnością (1) wartość bezwzględną z różnicy pojemność i jej wartości nominalnej. Zmiana w czasie parametru „*z*” została przedstawiona na rysunku 3.



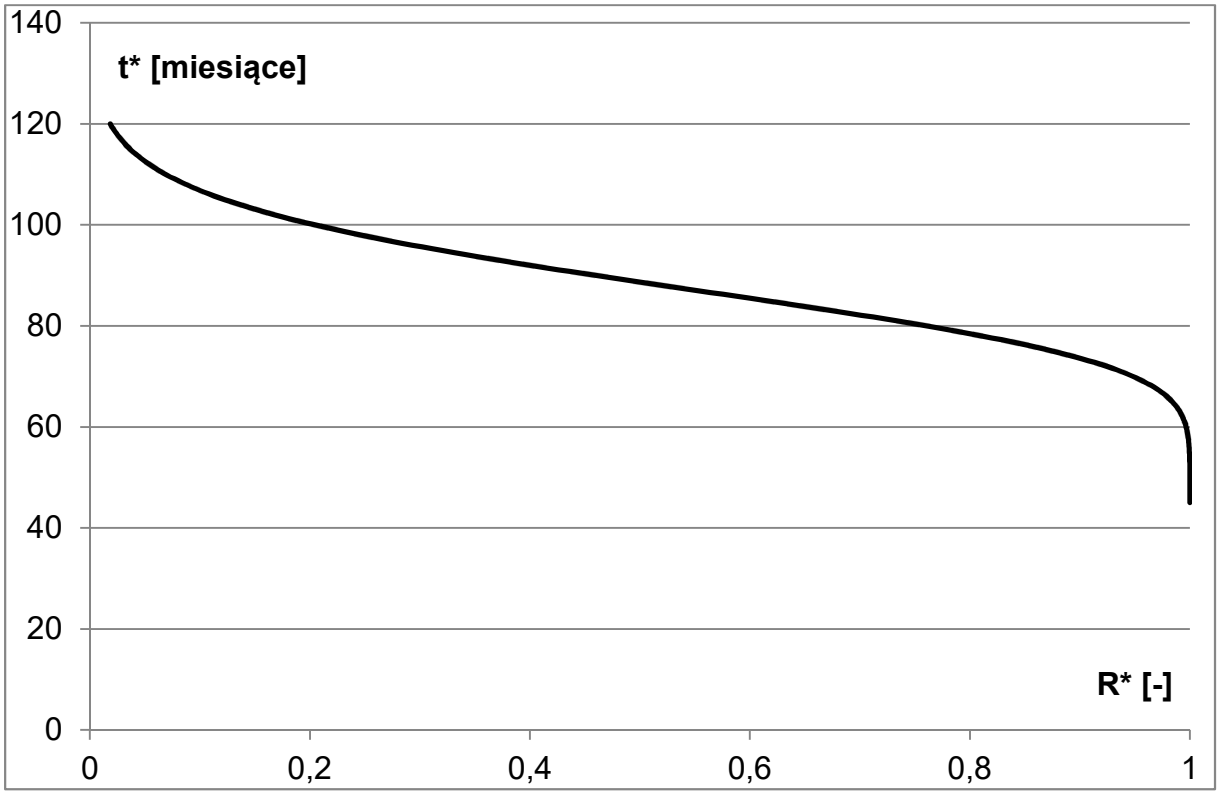
Rys. 3. Zmiana w czasie parametru „z” dla baterii 12-SAM-28

Posiadając zatem dane opisujące wartości parametru diagnostycznego w postaci $[(z_0,t_0),(z_1,t_1),(z_2,t_2),...,(z_n,t_n)]$, w oparciu o wzory (36) i (37) wyznaczono wartości współczynników funkcji gęstości:

$$b^* = 0{,}09, \qquad a^* = 0{,}015. \tag{38}$$

Parametr z_d określono posiłkując się dokumentacją techniczną wykorzystywaną do realizacji prac obsługowych, w której zamieszczono informację o wartości dopuszczalnej pojemności akumulatorów.

Dysponując zatem wartościami parametrów b_ε^*, a_ε^*, z_d podstawiono je do równań (11) lub (32) wyznaczając zależność czasu t^* od prawdopodobieństwa R^* - rysunek 4. W obu przypadkach (zależność (11) lub (32)) uzyskano ten sam przebieg.



Rys. 4. Zależność prognozowanej trwałości t^* od niezawodności R^*

Zakładając minimalną wartość niezawodności R* = 0,99 wyznaczono czas, do którego z założonym prawdopodobieństwem odchyłka parametru diagnostycznego nie przekroczy stanu granicznego:

$$T=63 \text{ [miesiące]} \tag{39}$$

Otrzymaną wartość (39) można wykorzystać w obsłudze technicznej w zależności od przyjętej strategii obsługi. W oparciu o powyższą metodykę można wyznaczyć kolejne okresy w których należy przeprowadzić kontrolę parametru diagnostycznego urządzenia [5,10].

5. Uwagi końcowe

W niniejszym opracowaniu przedstawiono zarys metody szacowania trwałości elementów lub zespołów gdy prędkość narastania zmian miała charakter losowy. Jednak sposób tej zmiany był opisany następującą, prostą zależnością

$$\frac{dz}{dt} = c\,,$$

gdzie c była zmienną losową określającą możliwość wzrostu odchyłki parametru.

Istnieje możliwość uogólnienia tej metody gdy prędkość narastania odchyłki będzie opisywana następującymi zależnościami:

$$\frac{dz}{dt} = cz\,, \tag{40}$$

$$\frac{dz}{dt} = ct^{\alpha-1}. \tag{41}$$

W pierwszym przypadku prędkość narastania zmian będzie miała charakter losowy zbliżony do wykładniczego. W drugim przypadku charakter narastania zmian będzie zbliżony do intensywności uszkodzeń w rozkładzie Weibulla.

Reasumując, można stwierdzić, że przedstawiona metoda wydaje się być poprawna i słuszna oraz pozwala na przeprowadzenie analizy stanu technicznego urządzenia ze względu na charakter zmian wartości parametrów diagnostycznych. Przedstawiony przykład obliczeniowy pozwolił na przeprowadzenie weryfikacji opracowanego modelu oraz uwypuklił aplikacyjne walory opracowanej metody. Metoda ta może być przydatna w dalszych pracach nad usprawnieniem zarówno procesu eksploatacyjnego jak i sposobu wykorzystania statków powietrznych z użyciem jego systemów pokładowych, pozwalając na określenie czasu przebywania urządzenia w stanie zdatności.

Ponadto z uwagi na swój uniwersalny charakter przedstawiona metoda, może z powodzeniem być zastosowana do określania trwałości resztkowej dowolnego obiektu technicznego, którego stan techniczny określany jest na podstawie analizy wartości parametrów diagnostycznych.

Przedstawiona metoda w niniejszym artykule może być dalej doskonalona i rozszerzona na inne przypadki narastania losowych zmian typu wykładniczego. Wydaje się, że może być wykorzystana do oceny niezawodności elementów mechanicznych, w przypadku rozpatrywania propagacji pęknięć zmęczeniowych w elementach poddanych losowemu obciążeniu i zastosowaniu wzoru Parisa do określenia prędkości pękania.

LITERATURA

1. DeLurgio SA. Forecasting principles and applications. University of Missouri-Kansas City: Irwin/McGraw-Hill, 1998.

2. Franck TD. Nonlinear Fokker-Planck Equations. Fundamentals and Applications. Berlin Heildenberg: Springer-Verlag, 2005.

3. Grasman J, Herwaarden OA. Asymptotic Methods for the Fokker-Planck Equation and the Exit Problem in Applications. Berlin Heildenberg: Springer-Verlag, 1999.

4. Idziaszek Z, Grzesik N. Object characteristics deterioration effect on task realizability – outline method of estimation and prognosis. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (3): 433–440.

5. Kinnison H, Siddiqui T. Aviation Maintenance Management. The McGraw-Hill Companies, Inc. 2013.

6. Knopik L, Migawa K. Multi-state model of maintenance policy. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (1): 125–130.

7. Knopik L, Migawa K, Wdzięczny A. Profit optimization in maintenance system, Polish Martime Research, 2016, 1(89): 193-98.

8. Kołowrocki K, Soszyńska Budny J. Reliability and Safety of Complex Technical Systems and Processes. Springer 2011.

9. McPherson JW. Reliability physics and engineering. New York: Springer, 2010.

10. Narayan V. Effective Maintenance Management. New York: Industrial Press Inc., 2012.

11. Pham H. Handbook of Engineering Statistics. London: Springer-Verlag 2006.

12. Rasuo B., Duknic G. Optimization of the aircraft general overhaul process. Aircraft engineering and aerospace technology 2013; 85 (5): 343-354.

13. Restel F. The Markov reliability and safety model of the railway transportation system. Safety and reliability: methodology and applications: proceedings of the European Safety and Reliability Conference, ESREL 2014, 14-18 September, 2015, Wrocław, Poland. CRC Press/Balkema: 303-311.

14. Risken H. The Fokker-Planck Equation. Methods of Solution and Applications. Berlin Heildenberg: Springer Verlag, 1984.

15. Tan CM, Singh P. Time evolution degradation physics in high power white LEDs under high temperature-humidity conditions. IEEE Transactions on Device and Materials Reliability 2014; 14(2): 742-750.

16. Ułanowicz L. Modelling of a process, which causes adhesive seizing (tacking) in precise pairs of hydraulic control devices. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18 (4): 492-500.

17. Valis D, Koucky M, Zak L. On approaches for non-direct determination of system deterioration. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 1:33-41.

18. Wang P, Tang Y, Baeb SJ, He Y. Bayesian analysis of two-phase degradation data based on change-point Wiener process. Reliability Engineering & System Safety 2018; 170: 244-256.

19. Wang YS, Zhang CH, Zhang SF, Chen X, Tan YY. Optimal design of constant stress accelerated degradation test plan with multiple stresses and multiple degradation measures. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2015; 229(1): 83-93.

20. Woch M. Reliability analysis of the PZL-130 Orlik TC-II aircraft structural component under real operating conditions. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (2): 287–295.

21. Zurek J, Tomaszek H. Zieja M. Analysis of structural component's lifetime distribution considered from the aspect of the wearing with the characteristic function applied. Safety, reliability and risk analysis: Beyond the horizon. Amsterdam: Balkema 2014, 2597-2602.

dr 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 25,
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 25,
email: szybja@agh.edu.pl

Metodyka szacowania niezawodności układów sieciowych na przykładzie komunikacji miejskiej

Słowa kluczowe: sieć komunikacyjna, niezawodność, algorytm faktoryzacji

Streszczenie: W artykule zaprezentowano opracowaną metodykę szacowania niezawodności układów sieciowych. Rozwiązanie to umożliwia dokonywanie oceny niezawodności oraz projektowanie modernizacji rozpatrywanej sieci przede wszystkim w aspekcie zapewnienia jej wymaganej niezawodności. Praktyczne wykorzystanie omawianej metodyki może mieć miejsce w odniesieniu do różnych układów sieciowych, np. sieci komputerowych, energetycznych, gazowych, wodociągowych, telekomunikacyjnych i transportowych. W artykule przedstawiono analizę niezawodności sieci komunikacyjnej w miejskim transporcie zbiorowym. Obliczenia przeprowadzono dla wybranych kryteriów zdatności sieci, które praktycznie warunkują jakość świadczonych usług transportowych. Podstawowe narzędzie obliczeniowe stanowił algorytm faktoryzacji, który umożliwia ocenę wpływu uszkodzeń poszczególnych połączeń (spowodowanych w szczególności czynnikami fizycznymi) na niezawodność całej sieci. W opracowaniu uwzględniono możliwość modernizacji analizowanej sieci, a uzyskane wyniki przedstawiono na wykresach.

1. Wprowadzenie

Opracowanie dotyczy problematyki szacowania niezawodności układów sieciowych, której ocena, ze względu na specyfikę tych układów, nie może być przeprowadzana tradycyjnymi metodami stosowanymi w teorii niezawodności. W artykule scharakteryzowano

tego typu układy, opracowano metodykę postępowania przy przeprowadzaniu obliczeń niezawodnościowych układów sieciowych oraz na wybranym przykładzie komunikacji miejskiej wykazano użyteczność zastosowanej metody faktoryzacji.

Działalność miejskich przedsiębiorstw transportowych jest silnie związana z gospodarowaniem majątkiem trwałym, którego składnikami w znacznej mierze są środki transportowe oraz infrastruktura w postaci szlaków i sieci komunikacyjnych. Istotnym elementem eksploatacji środków transportu jest gospodarka konserwacyjno-remontowa, będąca nieodzownym warunkiem zapewnienia ciągłości funkcjonowania i wykorzystania posiadanego potencjału majątkowego [2, 6, 27].

Gospodarka ta obejmuje całokształt działań zmierzających do utrzymania zdatności eksploatacyjnej sieci komunikacyjnych. Działania te wiążą się z koniecznością czasowego wyłączania z ruchu niektórych odcinków (linii) i wpływają negatywnie na wyniki działalności przedsiębiorstwa. Celem przedsiębiorstw jest takie kształtowanie struktury połączeń aby w sytuacji występujących awarii lub przeprowadzania planowych remontów linii trakcyjnych system komunikacji zapewnił użytkownikom osiąganie celu podróży, por. [10, 14, 31].

Sprostanie coraz wyższym wymaganiom w tym zakresie narzuca konieczność poszukiwania metod kształtowania i oceny systemów komunikacyjnych z uwzględnieniem sytuacji awaryjnych i kryteriów ekonomicznych. W tym zakresie wykorzystuje się rachunek prawdopodobieństwa (por. [11, 27, 33]). Na przykład analizuje się probabilistyczny wskaźnik nieregularności przejazdu zdefiniowany jako różnica (w minutach) między rzeczywistym a planowanym czasem w warunkach zapewnienia stabilnego i regularnego ruchu. Na wartość tego wskaźnika ma wpływ ukształtowanie sieci komunikacyjnej. Wyznaczenie jej korzystnej konfiguracji ma znaczenie dla wykonania zadania transportowego, por. [9, 18].

Do rozwiązania tego zadania można zastosować zaproponowany w pracy algorytm faktoryzacji [13, 24, 30]. Umożliwia on wyznaczenie wartości niezawodności dla złożonych układów sieciowych występujących w systemach transportu komunikacji miejskiej oraz stanowi alternatywę dla stosowanych w takich przypadkach mało dokładnych metod symulacyjnych [17, 19, 20].

2. Charakterystyka układów sieciowych

Do układów sieciowych zalicza się systemy: zaopatrzenia w energię, sieci teleinformatycznych, zaopatrzenia w wodę, transportowe, zapewniające ciągłość działania przedsiębiorstw produkcyjnych i dystrybucji produktów oraz inne, podobnego typu układy.

Wszystkie ww. należą do infrastruktury technicznej, która obejmuje urządzenia, sieci przesyłowe i związane z nimi obiekty techniczne realizujące zadania w zakresie dostarczania ciepła, gazu, wody, usuwania ścieków i odpadów, transportu i tym podobnych. Infrastruktura to podstawowe urządzenia i instytucje usługowe niezbędne do funkcjonowania gospodarki i społeczeństwa.

Infrastruktura liniowa w układzie sieciowym – charakteryzuje się liniowym połączeniem środków (elementów) technicznych realizujących określone funkcje.

Do oceny układów sieciowych można zastosować metody oparte o schematy blokowe. Stosowane mogą być do struktur niezawodnościowych sieciowych systemów o budowie modułowej i mogą mieć zastosowanie do dużych systemów transportowych w aglomeracjach

miejskich, które funkcjonują na znacznym obszarze oraz złożone są z wielu węzłów zawierających odwzorowane podsystemy wewnętrzne [12].

Modelowanie struktur sieciowych prowadzone jest często z użyciem grafów składających się z węzłów oraz odcinków łączących węzły. Obliczenia wartości niezawodności dla struktur odwzorowanych w postaci grafów mogą być efektywnie realizowane z zastosowaniem metod faktoryzacji [17, 19, 23]. W systemach transportu komunikacji miejskiej struktury niezawodnościowe cały czas mogą ewoluować. Przyczyną takiego stanu może być zmiana obszaru objętego działaniem, albo np. zmieniający się stan zapotrzebowania na usługi transportowe. Poprawna identyfikacja struktur rzeczywistych systemów technicznych oraz metoda wyliczenia wartości wskaźników niezawodności jest zabiegiem kluczowym do prawidłowej analizy niezawodności i oceny zagrożenia dla efektywnego funkcjonowania systemu komunikacji miejskiej.

Trudności w ocenie niezawodności układów sieciowych wynikają z faktu, że nie posiadają uznanej za klasyczną struktury niezawodnościowej, mogą występować zależności uszkodzeń między fragmentami układu oraz ograniczone jest zastosowanie typowych wskaźników i charakterystyk niezawodnościowych. Z tego względu do oceny niezawodności tego typu układów można zastosować algorytm faktoryzacji, wykorzystujący teorię grafów i metody redukcji sieci [4, 23, 24]. Istotnym aspektem zapewnienia niezawodności eksploatacyjnej układu o strukturze sieci jest kształtowanie i optymalizowanie jego struktury połączeń już na etapie projektowania [5, 13, 26]. Poprawa wiarygodności dokonywanej oceny niezawodności sieci możliwa jest natomiast poprzez uwzględnienie uszkadzania się i odnawiania w procesie eksploatacji wszystkich jej elementów [25].

3. Metodyka badań niezawodnościowych układów sieciowych

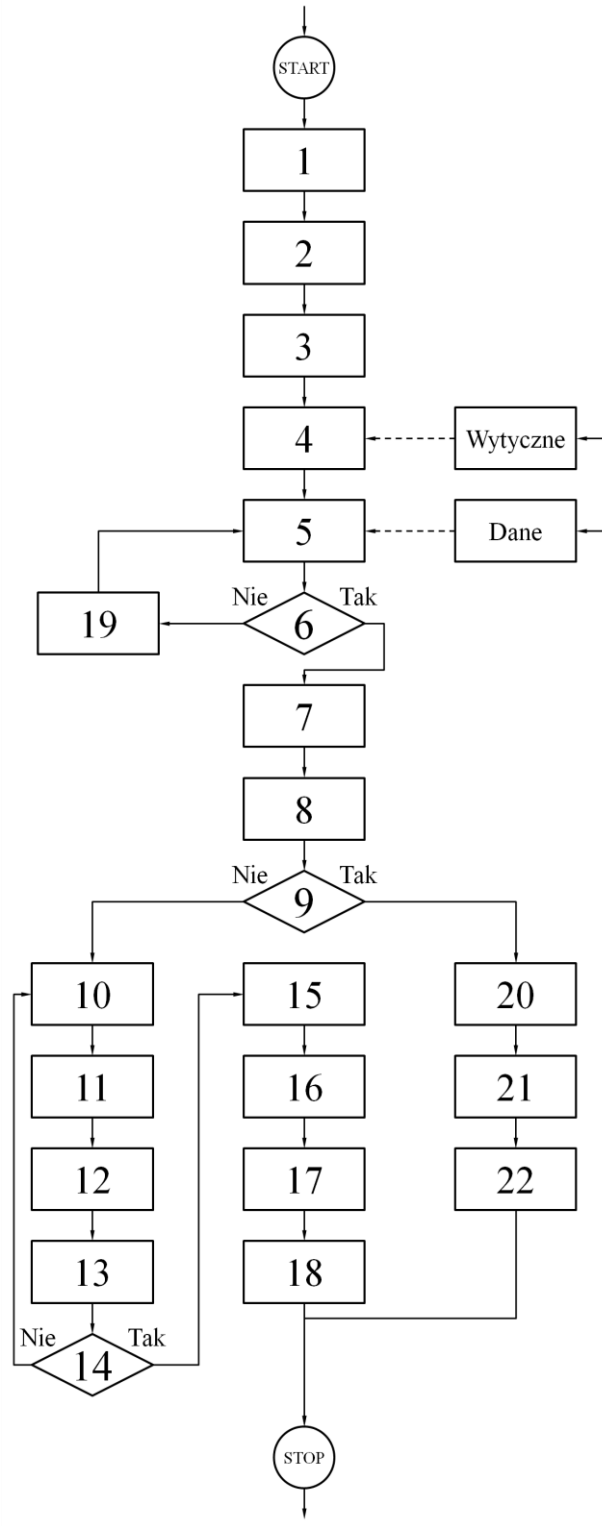
Ocena niezawodności układów sieciowych wymaga zastosowania określonych modeli matematycznych oraz metod związanych z ich wykorzystaniem. Działanie takie wymaga jednak wcześniejszego przeprowadzenia odpowiednich prac przygotowawczych oraz późniejszego stosownego opracowania uzyskanych wyników [30]. Ogół działań, które towarzyszyć mogą ocenie niezawodności układu sieciowego, został zestawiony i uporządkowany w ramach opracowanej metody oceny i modernizacji sieci. Jej schemat przedstawiono na rysunku 1.

Doraźnym celem analizy układu sieciowego może być wyznaczenie wskaźników charakteryzujących jego niezawodność, jednak takie działanie powinno jedynie wspomagać inne, bardziej złożone działania o donioślejszych skutkach. Należą do nich zwłaszcza: planowanie odnawiania profilaktycznego rozpatrywanej sieci oraz przygotowanie modernizacji sieci (szczególnie na drodze zmiany jej struktury niezawodnościowej).

Opracowana metoda nie uwzględnia działań związanych z planowaniem odnawiania profilaktycznego (co nie stanowi przedmiotu zainteresowania w ramach niniejszego opracowania), jednak pokazuje ona, że działania o charakterze modernizacji powinny zostać poprzedzone rzetelną oceną niezawodności rozpatrywanego układu sieciowego.

Dokonywana ocena niezawodności wymaga wcześniejszych działań przygotowawczych (co widoczne jest na schemacie pokazanym na rysunku 1) jak również dostarczenia precyzyjnie sformułowanych wymagań odnośnie do poprawnego

funkcjonowania analizowanego układu sieciowego oraz danych na temat przebiegu jego dotychczasowej eksploatacji. W sytuacji, gdy jakość dostępnych danych nie jest zadowalająca, wykorzystywane mogą być zastępcze źródła danych nt. przebiegu procesu eksploatacji, w postaci: zapisów księgowych, zapisów ewidencji wykorzystania części wymiennych oraz materiałów, itp. (por. [30]).



1. Sformułować cel prowadzonej analizy.
2. Zamodelować rozpatrywaną sieć jako układ węzłów i połączeń.
3. Nazwać poszczególne węzły i połączenia.
4. Określić funkcję badanej sieci i kryterium jej zdatności.
5. Zweryfikować dane nt. niezawodności sieci.
6. Czy jakość dostępnych danych nt. niezawodności sieci jest dostateczna?
7. Opracować dane nt. poszczególnych węzłów i połączeń.
8. Ocenić niezawodność sieci zgodnie z kryterium (por. pkt. 4), z zastosowaniem modeli matematycznych.
9. Czy celem prowadzonej analizy (por. pkt 1) jest wyłącznie ocena niezawodności sieci?
10. Wytypować „słabe ogniwa" badanej sieci.
11. Opracować wariant modernizacji rozpatrywanej sieci.
12. Przyjąć prognozowaną niezawodność nowych odcinków sieci.
13. Oszacować niezawodność $R_P(t)$ zmodernizowanej sieci.
14. Czy $R_P(t) \geq R_W(t)$?
15. Sprawdzić całość rozwiązania.
16. Opracować stosowną dokumentację techniczną.
17. Zweryfikować poprawność dokumentacji.
18. Zatwierdzić całość rozwiązania i dokumentacji.
19. Zaplanować i przeprowadzić stosowne badania eksploatacyjne lub wykorzystać zastępcze źródła danych.
20. Opracować wyniki analizy.
21. Zweryfikować poprawność przeprowadzonej analizy.
22. Zapisać wnioski z przeprowadzonej analizy.

Rys. 1. Schemat metody oceny niezawodności i modernizacji układu sieciowego

Poszczególne działania i zapytania przedstawione na schemacie za pomocą bloków z odpowiednimi liczbami (kroki od 1 do 22) zostały opisane na rysunku. Przedstawiony schemat blokowy opisuje procedurę postępowania, w nawiązaniu do której realizowano analizę niezawodnościową sieci tramwajowej scharakteryzowanej w rozdziale 4. Opisy w poszczególnych blokach pokazanych na rysunku 1 mają charakter ogólny, co jest konsekwencją uniwersalności opracowanej metody, która może być wykorzystywana podczas różnego rodzaju analiz i prac projektowych dotyczących niezawodności układów sieciowych.

W opracowanej metodzie przyjęto, że modernizacja sieci dokonywana będzie na drodze wprowadzania kolejnych zmian w jej strukturze niezawodnościowej do chwili, gdy jej szacowana niezawodność $R_P(t)$ stanie się niemniejsza od wymaganej $R_W(t)$ – dokonywane to jest w pętli iteracyjnej obejmującej kroki od 10 do 14. Tego rodzaju metoda kształtowania niezawodności została przedstawiona (w odniesieniu do obiektów technicznych innej klasy) w opracowaniach [7, 8], por. [29].

4. Przykład przeprowadzenia analizy niezawodnościowej sieci komunikacyjnej w miejskim transporcie zbiorowym

Wymagania dotyczące jakości usług w komunikacji miejskiej w wysokim stopniu uzależnione są od systemu organizacji, od technicznych środków transportu oraz struktury połączeń sieci transportowej i możliwości jej wykorzystania. Decydujące o jakości i niezawodności w miejskim transporcie zbiorowym są kryteria, mające istotne znaczenie dla użytkownika transportu zbiorowego oraz eksploatatora sieci, który uwzględniać musi również czynniki związane z ekonomiką eksploatacji.

Poddając analizie różne kryteria jakościowe i ekonomiczne w komunikacji miejskiej, do strategicznych zaliczyć można te, które związane są z infrastrukturą sieci połączeń oraz pojazdami realizującymi zadania przewozowe. Do najważniejszych z tych kryteriów zaliczyć można [1, 3, 16, 22]:

- dostępność linii wynikająca z liczby i rozmieszczenia przystanków,
- integrację przestrzenną i funkcjonalną z innymi środkami transportu zbiorowego (krajowego, regionalnego i lokalnego, *Park and Ride, Bike and Ride*, itp.),
- bezpośredniość połączeń i elastyczność funkcjonowania rozumianą jako możliwość wyboru połączenia alternatywnego,
- niezawodność funkcjonowania, której kryteriami są: punktualność, regularność, prawdopodobieństwo dojazdu do celu, prawdopodobieństwo osiągnięcia celu podróży w spodziewanym czasie lub osiągnięcie celu końcowego trasy.

Struktura sieci komunikacyjnej determinuje możliwości połączeń między węzłami, a te wpływają bezpośrednio na możliwości transportowe całego układu i na kształtowanie opinii użytkowników o jakości systemu komunikacji. Pod pojęciem węzła w strukturze sieci komunikacyjnej przyjęto miejsca skrzyżowania i rozjazdów torów oraz przystanki umożliwiające przesiadkę na inny rodzaj środka transportu, a także przystanki początkowe i końcowe. Racjonalnie ukształtowana infrastruktura w postaci efektywnej lokalizacji strategicznych węzłów komunikacyjnych daje możliwości optymalizacji sieci połączeń i

związanej z tym wysokiej niezawodności systemu. Natomiast istotne znaczenie dla efektywności i jakości całego systemu ma również poziom niezawodności węzłów i odcinków komunikacyjnych ze szczególnym wskazaniem na węzły o znaczeniu strategicznym.

Strukturę połączeń komunikacyjnych aglomeracji przedstawić można w postaci grafu wyrażającego zależności funkcjonalne pomiędzy elementami tego systemu i analizować je w ujęciu niezawodnościowym. Można go interpretować jako sieć połączeń w systemach transportu szynowego komunikacji miejskiej między węzłami infrastruktury. Realizacja procesu transportowego odbywa się między węzłami początkowymi (zajezdnie) i węzłami końcowymi (pętle końcowe), kolejne elementy są węzłami systemu jako cząstkowe realizacje zadań (przystanki), natomiast linie są infrastrukturą szynową między węzłami. Przy takiej interpretacji, zdatność lub niezdatność odcinka odpowiada za wykonanie zadania przewozowego przez jedną lub więcej linii albo niewykonanie zadania przewozowego przez przyporządkowaną (którąkolwiek) linię od początku do końca. Założono, że zdatności systemu odpowiada wykonanie zadania przez wszystkie linie, natomiast niezdatność jest interpretowana jako brak wykonania zadania przewozowego już przez jedną linię [12, 19].

W pracy zaproponowano zastosowanie algorytmu faktoryzacji w analizie strukturalnej sieci do oceny niezawodności połączeń komunikacyjnych w miejskim transporcie zbiorowym.

Metoda faktoryzacji oparta jest na następujących założeniach [4, 21, 25, 32]:
- analizowaną sieć reprezentuje graf nieskierowany $G=(V, E)$, w którym $V=\{v_1, v_2,\ldots, v_n\}$ reprezentuje zbiór wierzchołków grafu (węzły w sieci) a $E=\{e_1, e_2,\ldots, e_m\}$ reprezentuje zbiór krawędzi grafu (połączenia w sieci),
- wszystkie połączenia e_i w sieci ulegają wzajemnie niezależnym uszkodzeniom ze znanym prawdopodobieństwem,
- niezawodność sieci R_s jest przyjmowana jako prawdopodobieństwo połączenia między wybranym w sieci zbiorem węzłów K, którego liczność może wynosić: $2 \le |K| \le |V|$,
- wybór zbioru węzłów K określa wyznaczaną miarę niezawodności sieci (K – terminal reliability network)
- proces wyznaczania zależności do obliczenia niezawodności sieci oparty jest na zasadzie ściągania i wycinania połączeń, którą stosuje się cyklicznie do wszystkich krawędzi grafu reprezentującego sieć, co zapisuje się w postaci formuły:

$$R_s = R(G_K) = R_{e_i} \cdot R(G_{K'} * e_i) + (1 - R_{e_i}) \cdot R(G_K - e_i) \tag{1}$$

gdzie:
e_i – i-te połączenie w sieci reprezentowanej przez graf G,
R_{ei} – prawdopodobieństwo, że połączenie e_i jest zdatne,

$$G_{K'} * e_i = (V - v_k - v_l + v_{kl}, E - e_i), \quad v_{kl} = v_k \cup v_l,$$

$$K' = \begin{cases} K & \text{if} \quad v_k \quad or \quad v_l \notin K, \\ K - v_k - v_l + v_{kl} & \text{if} \quad v_k \in K \quad and \quad v_l \in K, \end{cases}$$

$$G_K - e_i = (V, E - e_i),$$

- po zredukowaniu wszystkich połączeń otrzymuje się zależność będącą analitycznym zapisem struktury niezawodnościowej, umożliwiającym obliczenie wybranej miary niezawodności sieci.

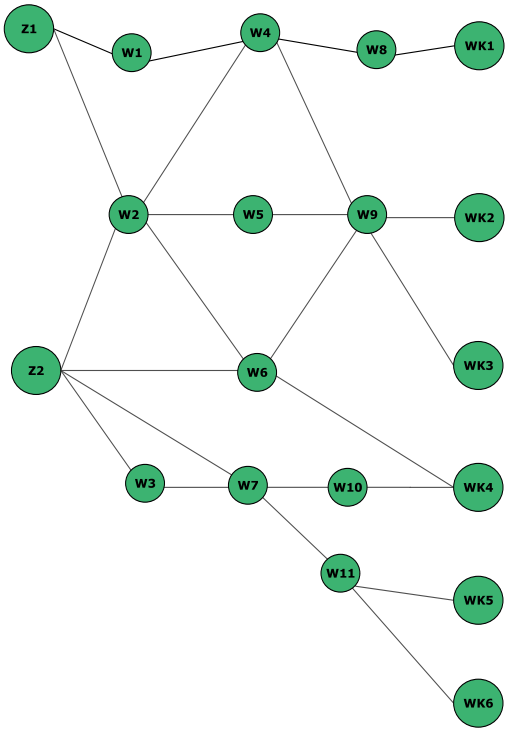
Metoda faktoryzacji umożliwia obliczenia niezawodności układów sieciowych, co nie jest możliwe do zrealizowania z zastosowaniem klasycznych metod oceny niezawodności. Efektem tych obliczeń jest oczywiście uzyskanie typowych charakterystyk niezawodności całego układu, jednak droga prowadząca do ich wyznaczenia nie mieści się w ramach standardowego postępowania przy klasycznym podejściu do szacowania niezawodności obiektów technicznych.

Systemy komunikacji miejskiej mogą być odwzorowywane jako struktury sieciowe dla których trzeba wyróżnić wiele wejść z i wyjść w_k natomiast w ich opisie grafowym mamy wtedy do czynienia z grafami lub sieciami takimi jak przedstawiona na rysunku 1.Jako warunek zdatności takiego grafu zazwyczaj przyjmuje się istnienie połączenia od każdego węzła początkowego (wejściowego) $z \in \boldsymbol{Z}$ do dowolnego węzła końcowego (wyjściowego) $w \in \boldsymbol{W_k}$, przy czym pewne węzły grafu, mogą być zarówno węzłami wejściowymi, jak i wyjściowymi (tzn. $\boldsymbol{Z} \cap \boldsymbol{W_k} \neq \varnothing$). W analizowanym w pracy przypadku podjęto założenia, że wykonanie zadania będzie zrealizowane w przypadku gdy z wyznaczonego węzła $\boldsymbol{z_1}$pojazdy dojadą do wyznaczonych odpowiednich węzłów od $\boldsymbol{w_{k1}}$ do $\boldsymbol{w_{k4}}$ i podobnie z $\boldsymbol{z_2}$ do odpowiednich od $\boldsymbol{w_{k2}}$ do $\boldsymbol{w_{k6}}$.

Dla potrzeb badań i przedstawienia możliwości obliczeniowych w opracowaniu zaproponowano model struktury połączeń sieci w postaci grafu nieskierowanego, co przedstawiono na rysunku 2. Do analizy przyjęte zostały szacunkowe wartości wskaźników niezawodności połączeń tej struktury przy założeniu zachowania zdatności węzłów.

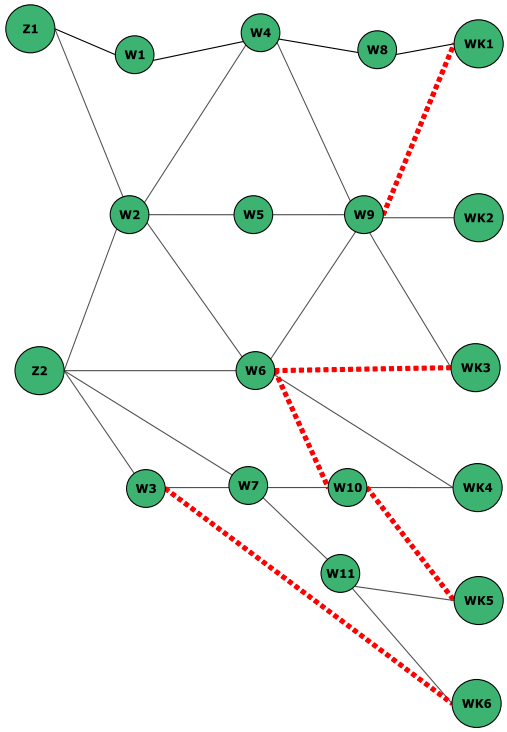
Wartości tych charakterystyk niezawodnościowych odzwierciedlają skutki wszelkiego rodzaju oddziaływań zewnętrznych na badany układ. Czynniki zewnętrzne mogą mieć charakter losowy (przypadkowy) albo mogą podlegać sterowaniu przez eksploatatorów układów. Ocena wartości charakterystyk ma charakter statystyczny i jest wynikiem obserwacji rzeczywistych systemów eksploatacji. W związku z tym, w artykule nie systematyzowano utrudnień występujących w obrębie sieci transportowej w obszarze miejskim, traktując przyjęte charakterystyki jako kompleksową reprezentację wszystkich zewnętrznych oddziaływań.

Następnie wprowadzono zmiany w strukturze sieci przez zwiększenie liczby połączeń miedzy węzłami zaznaczone liniami kropkowymi. Schemat struktury po modyfikacji przedstawiono na rysunku 3.



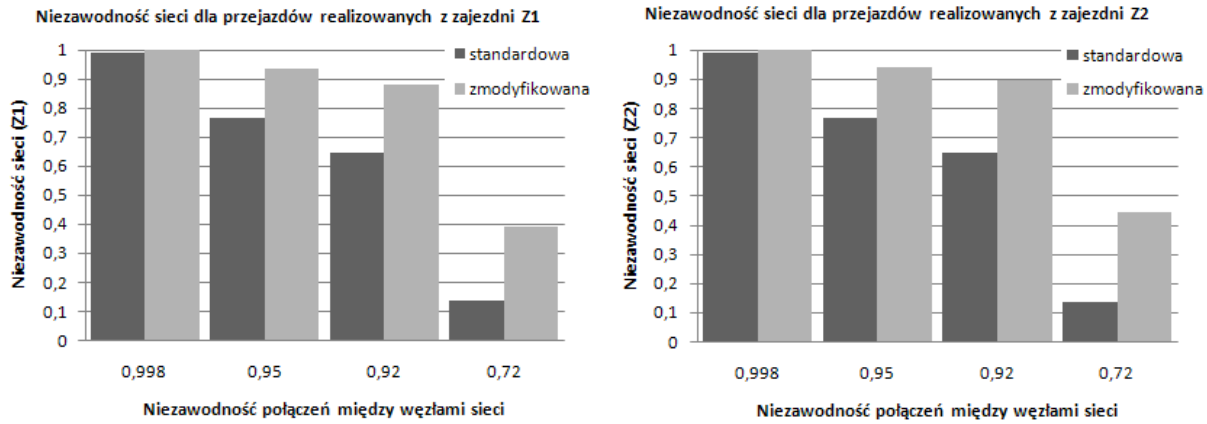
Rys. 2. Schemat struktury sieci komunikacji miejskiej

Dla zmodyfikowanej struktury ponownie przeprowadzono obliczenia i określono, powstałe w jej efekcie, zmiany niezawodności sieci. W przeprowadzonej analizie przyjęto, że miarą niezawodności sieci, wynikającą z algorytmu faktoryzacji, jest prawdopodobieństwo istnienia (występowania) połączenia między wybraną zajezdnią (Z1 lub Z2) a wszystkimi punktami końcowymi tras wychodzących z danej zajezdni (dla zajezdni Z1 węzłami końcowymi są WK1-WK4, a dla zajezdni Z2 węzły końcowe to WK2-WK6). Jest to inaczej prawdopodobieństwo realizacji (dotarcia do punktu docelowego) wszystkich planowanych przejazdów z danej zajezdni.



Rys. 3. Schemat struktury sieci komunikacji miejskiej po modyfikacji

Na rysunku 4 przedstawiono ww. miary niezawodności sieci obliczone dla sieci standardowej (Rys. 2) oraz zmodyfikowanej (Rys. 3), przy różnych wartościach niezawodności połączeń między węzłami: R=0,998; R=0,95; R=0,92 oraz R=0,72. Do przeprowadzenia obliczeń wykorzystano program komputerowy, napisany w języku C, który realizuje obliczenia zgodnie z przedstawioną procedurą algorytmu faktoryzacji.



Rys. 4. Niezawodność sieci dla przejazdów realizowanych z zajezdni Z1 i Z2

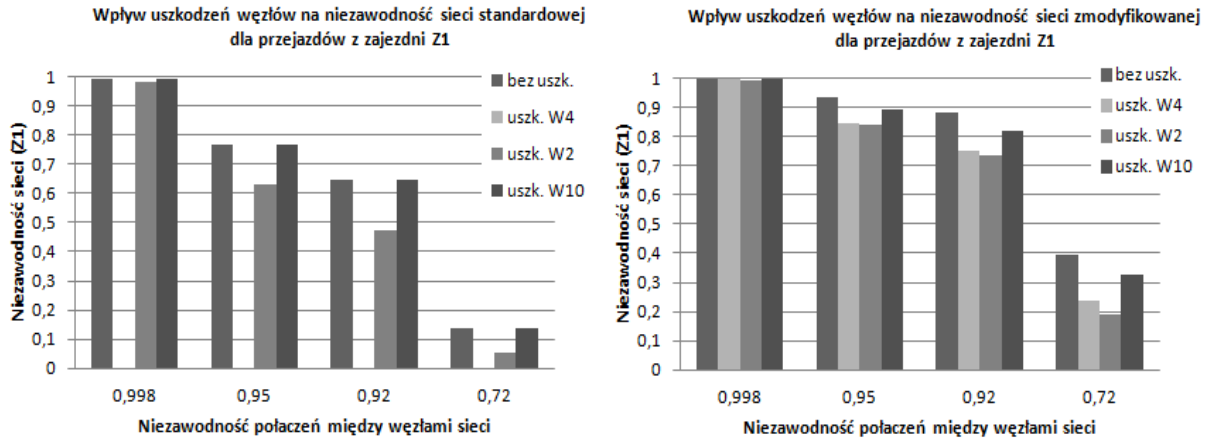
Można zauważyć, że tylko przy dużych wartościach niezawodności połączeń między węzłami (R=0,998) wprowadzone modyfikacje nie zmieniają znacząco niezawodności sieci (względna zmiana dla Z1 to 0,8% i dla Z2 to również 0,8%). W pozostałych przypadkach im mniejsza jest niezawodność połączeń tym większa jest zmiana niezawodności sieci w efekcie

jej modyfikacji. W przypadku skrajnym (R=0,72) względne zmiany wynoszą odpowiednio: dla Z1 65,2% a dla Z2 69,8%.

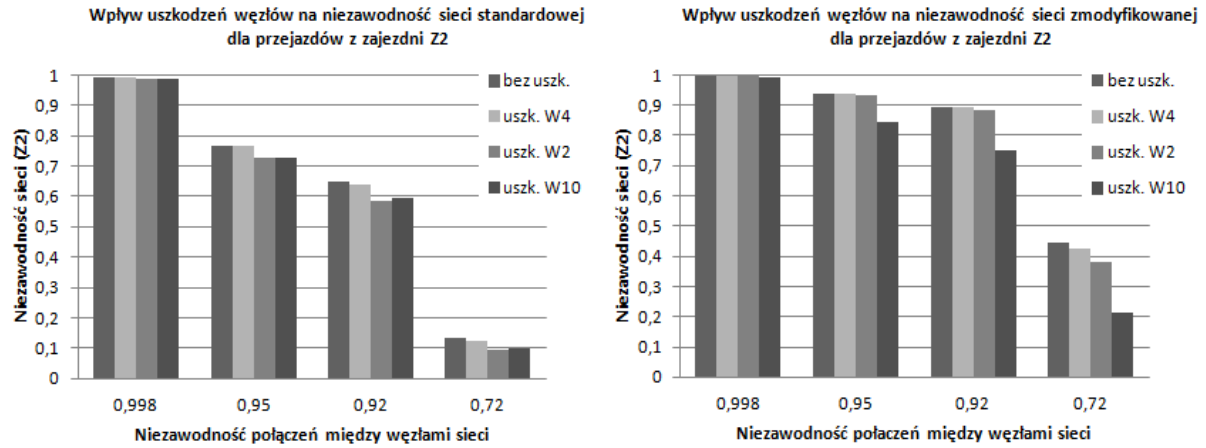
Dodatkowo przeprowadzono analizę niezawodności sieci z uwzględnieniem uszkodzeń węzłów. Uszkodzenie węzła traktowane jest jako uszkodzenie wszystkich połączeń skojarzonych z danym węzłem.

W związku z powyższym wyróżnić można tzw. węzły krytyczne, czyli węzły, których uszkodzenie powoduje brak możliwości realizacji przynajmniej jednego z planowanych przejazdów z danej zajezdni do punktów końcowych tras. W przypadku sieci standardowej (Rys. 2) są to węzły: W4, W8, W9, W6, W7 i W11. W przypadku sieci zmodyfikowanej (Rys. 3) znacznie zmniejsza się liczba węzłów krytycznych, którymi pozostają tylko węzły W8 i W6.

Biorąc pod uwagę uszkodzenia wybranych węzłów (W4, W2, W10), które oprócz W4 w analizowanej sieci nie są węzłami krytycznymi, dla przypadku sieci standardowej otrzymano wyniki przedstawione w postaci histogramów (Rys. 5 i 6).



Rys. 5. Wpływ uszkodzeń węzłów na niezawodność sieci dla przejazdów realizowanych z zajezdni Z1



Rys. 6. Wpływ uszkodzeń węzłów na niezawodność sieci dla przejazdów realizowanych z zajezdni Z2

Skrajne wartości w formie względnych różnic procentowych przedstawiają się następująco:

- dla Z1 uszkodzenie węzła W4 zmniejsza niezawodność o 100% we wszystkich przypadkach (jest to węzeł krytyczny),
- dla Z2 uszkodzenie węzła W4 zmniejsza niezawodność w skrajnych przypadkach o ≈0% dla R=0,998 oraz o 7,7% dla R=0,72,
- dla Z1 uszkodzenie węzła W2 zmniejsza niezawodność w skrajnych przypadkach o 0,8% dla R=0,998 oraz o 61,9% dla R=0,72,
- dla Z2 uszkodzenie węzła W2 zmniejsza niezawodność w skrajnych przypadkach o 0,2% dla R=0,998 oraz o 30,7% dla R=0,72,
- dla Z1 uszkodzenie węzła W10 nie zmienia niezawodności,
- dla Z2 uszkodzenie węzła W10 zmniejsza niezawodność w skrajnych przypadkach o 0,2% dla R=0,998 oraz o 25,5% dla R=0,72.

W sposób analogiczny przedstawiono wyniki obliczeń dla sieci zmodyfikowanej (Rys. 4 i 5):

- dla Z1 uszkodzenie węzła W4 zmniejsza niezawodność w skrajnych przypadkach o 0,4% dla R=0,998 oraz o 38,9% dla R=0,72,
- dla Z2 uszkodzenie węzła W4 zmniejsza niezawodność w skrajnych przypadkach o ≈0% dla R=0,998 oraz o 3,7% dla R=0,72,
- dla Z1 uszkodzenie węzła W2 zmniejsza niezawodność w skrajnych przypadkach o 0,4% dla R=0,998 oraz o 51,6% dla R=0,72,
- dla Z2 uszkodzenie węzła W2 mniejsza niezawodność w skrajnych przypadkach o ≈0% dla R=0,998 oraz o 14,5% dla R=0,72
- dla Z1 uszkodzenie węzła W10 zmniejsza niezawodność w skrajnych przypadkach o 0,2% dla R=0,998 oraz o 16,8% dla R=0,72,
- dla Z2 uszkodzenie węzła W10 zmniejsza niezawodność w skrajnych przypadkach o 0,4% dla R=0,998 oraz o 51,5% dla R=0,72.

5. Podsumowanie

Ocena niezawodności układów sieciowych wymaga zastosowania wyspecjalizowanych metod obliczeniowych, które nie są prezentowane w klasycznej teorii niezawodności.

Do tych metod można zaliczyć algorytm faktoryzacji, który nadaje się do szybkiej oceny zmian efektywności rezerwowania w zależności od liczby i lokalizacji punktów zasilania, do doboru optymalnej liczby źródeł zasilania i ich usytuowania ze względu na zapewnienie wysokiej niezawodności zasilania badanego fragmentu sieci, do projektowania sieci z uwzględnieniem niezawodności i opłacalności (efektywności) rezerwowania. Zastosowanie jej pozwala na obserwację zmian niezawodności sieci przy wprowadzanych modyfikacjach tej struktury (przez np.: dodawanie kolejnych połączeń), co będzie użyteczne w praktyce dla projektantów nowych sieci oraz przy planowanych modernizacjach i rozbudowie istniejących sieci. Wykorzystany w powyższej analizie algorytm faktoryzacji umożliwia wyznaczenie niezawodności sieci o dowolnej konfiguracji.

W miejsce klasycznych wskaźników i charakterystyk niezawodnościowych proponowane jest zastosowanie parametru strumienia uszkodzeń do oceny niezawodności, co nie tylko nie ogranicza możliwości modernizacji badanych układów, ale zwiększa utylitarny charakter przeprowadzanych analiz niezawodnościowych.

Wielokrotne i szybkie powtarzanie obliczeń dla różnych wariantów modyfikacji sieci wymaga zapisania algorytmu w postaci programu komputerowego.

Ocena niezawodności sieci rozbudowana o analizę ekonomiczną oraz ryzyka związanego z eksploatacją może być użyteczna dla projektantów sieci komunikacyjnych, przy planowanych modernizacjach oraz rozbudowie istniejących sieci.

W artykule przedstawiono oceny niezawodności wybranych fragmentów sieci komunikacyjnej przyjmując, że możliwości zrealizowania przejazdów między wybranymi węzłami sieci stanowi istotne kryterium dla użytkowników (pasażerów) środków komunikacji miejskiej. Możliwości występowania awarii między połączeniami w węzłach sieci są obserwowane w rzeczywistych warunkach eksploatacji i stanowią poważne utrudnienie w realizacji zadań transportu szynowego komunikacji miejskiej. Z tego też względu przedstawiona analiza niezawodnościowa powinna być użyteczna dla miejskich przedsiębiorstw komunikacyjnych w planowaniu i rozbudowie układów trakcyjnych.

Metoda faktoryzacji jest przedstawiona w prezentowanej poniżej literaturze i nie charakteryzowano jej szczegółowo prezentując jedynie wyniki obliczeń ponieważ spowodowałoby to znaczne zwiększenie objętości artykułu, którego celem jest prezentacja metodyki szacowania niezawodności układów sieciowych, a nie metody faktoryzacji.

Literatura

1. Borowiecki R, Kaczmarek J, Magiera J, Młynarski S. Eksploatacja taboru szynowego komunikacji miejskiej. Niezawodność, jakość, ekonomika. Kraków: Wydawnictwo Akademii Ekonomicznej w Krakowie, 2004.
2. Brons M, Nijkamp P, Pels E, Rietveld R. Efficiency of urban public transit: A meta analysis. Transportation 2005; 32: 1–21.
3. Bryniarska Z. Ocena satysfakcji pasażerów z innowacyjnego rozwiązania w miejskim transporcie zbiorowym w Krakowie. Transport miejski i regionalny 2017; 11: 5–9.
4. Carlier J, Lucet C. A decomposition algorithm for network reliability evaluation. Discrete Applied Mathematics 1996; 65: 141–156.
5. Cats O, Koppenol GJ, Warnier M. Robustness assessment of link capacity reduction for complex networks: Application for public transport systems. Reliability Engineering and System Safety 2017; 167; DOI: 10.1016/j.ress.2017.07.009.
6. Crainic TG, Ricciardi N, Storchi G. Models for Evaluating and Planning City Logistics Systems. Transportation Science 2009; 43(4): 432–454.
7. Heinrich M, Lenkiewicz W. Erhöhung und Auswertung des Zuverlässigkeitsniveaus von tribologischen Paarungen im Zeitraum der Anfangsnutzung eines einzelnen Objekts. Tribologie und Schmierungstechnik 2002, 3.
8. Heinrich M. Badania eksploatacyjne jednostkowych złożonych obiektów technicznych w celu podniesienia ich niezawodności konstrukcyjnej. Zagadnienia Eksploatacji Maszyn 1993; 1-2 (93-94): 125–133.
9. Holmgren J. The efficiency of public transport operations – An evaluation using stochastic frontier analysis. Research in Transportation Economics 2013; 39 (1): 50–57.

10. Jan R-H. Design of reliable networks. Computers & Operations Research 1993; 20: 25–34.
11. Jong GDe, Cheung F. Stochastic frontier models for public transport. Meersman H, van de Voorde E, Winkelmans W eds. World Transport Research: Selected Proceedings of the 8th World Conference on Transport Research. Transport Modes and Systems. New York: Pergamon, 1999; 1: 373–386.
12. Karpiński J, Korczak E. Metody oceny niezawodności dwustanowych systemów technicznych. Omnitech Press, 1990.
13. Koide T, Shinmori S, Ishii H. Topological optimization with a network reliability constraint. Discrete Applied Mathematics 2001; 115: 135–149.
14. Levaggi R. Parametric and non-parametric approach to efficiency: The case of urban transport in Italy. Studi Economici 1994; 49: 67–88.
15. Lucet C, Manouvrier J-F. Exact methods to compute network reliability, in Statistical and Probabilistic Models in Reliability. Ionescu DC and Limnios N eds. Boston: Birkhauser, 1999: 279–294.
16. Malasek J. Metoda oceny dostępności i atrakcyjności przystanków miejskiego transportu zbiorowego. Transport miejski i regionalny 2017; 9: 26–32.
17. Młynarski S. Niezawodność strukturalna w systemach logistycznych ratownictwa. Logistyka [CD] 2014; 4: 871–880.
18. Nolan JF. Determinants of productive efficiency in urban transit. Logistics and Transportation Review 1996; 32: 319–342.
19. Oprzędkiewicz J, Młynarski S. Problems of economic safety and intelligence system and reliability of global systems. In Enterprises in the face of 21st century challenges. Development – management – entrepreneurship. Borowiecki R & Jaki A eds. Cracow: Cracow University of Economics, 2008: 307–315.
20. Oprzędkiewicz J. Komputerowa metoda oceny niezawodności systemów technicznych. Lublin: LTN, 1997.
21. Page LB, Perry JE. A practical implementation of the factoring theorem for network reliability. IEEE Transactions on Reliability 1988; 37: 259-267.
22. Paszkowski J, Kucharski R. Paradoksy przepustowości miejskiej sieci drogowej i sposoby ich odwzorowania w modelu czterostadiowym. Transport miejski i regionalny 2017; 10: 5–11.
23. Pilch R, Szybka J. Koncepcja zastosowania algorytmu faktoryzacji do oceny niezawodności ciągów komunikacyjnych – Application of factoring algorithms for estimation of road network reliability. Problemy Eksploatacji – Maintenance Problems 2007; 2: 129–136.
24. Pilch R. Factorisation algorithm-based method used for the calculation of network system's reliability – Metodyka wyznaczania niezawodności układów sieciowych w oparciu o algorytm faktoryzacji. Zagadnienia Eksploatacji Maszyn – Scientific Problems of Machines Operation and Maintenance 2011; 4 (168): 45–57.
25. Pilch R. Reliability evaluation of networks with imperfect and repairable links and nodes. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (1): 19–25.
26. Ramirez-Marquez JE, Rocco CM. All-terminal network reliability optimization via probabilistic solution discovery. Reliability Engineering and System Safety 2008; 93: 1689–1697.
27. Rymarz J, Niewczas A, Krzyżak A. Comparison of operational availability of public city buses by analysis of variance. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18 (3): 373–378, http://dx.doi.org/10.17531/ein.2016.3.8.
28. Smalko Z. Modelowanie eksploatacyjnych systemów transportowych. Radom: ITE, 1996.

29. Smolnik M. A conception of modernising LEMACH designing methods using TRIZ instruments. Souchkov V, Kässi T. eds. TRIZfest-2014. Theories and applications. The 10th international conference: September 4–6, 2014, Prague, Czech Republic. Conference proceedings. Knoxville: International TRIZ Association – MATRIZ, 2014.
30. Smolnik M. Koncepcja systemu informacyjnego wspomagającego kierowanie eksploatacją złożonych odnawialnych obiektów technicznych. Logistyka 2015; 5: 1307–1312.
31. Wiśniewski P. Projekt i implementacja systemu wspomagania decyzji dla zarządzania kryzysowego transportem miejskim. Transport miejski i regionalny 2017; 11: 10–16.
32. Wood RK. Factoring algorithms for computing K-terminal network reliability. IEEE Transactions on Reliability 1986; 35 (3): 269–278.
33. Zamojski W. Niezawodność i eksploatacja systemów. Wrocław: Politechnika Wrocławska, 1981.

Andrzej ŚWIDERSKI[1]
Arkadiusz JÓŹWIAK[2]
Roland JACHIMOWSKI[3]

[1]Instytut Transportu Samochodowego, Warszawa, ul. Jagiellońska 80, 03-301 Warszawa, andrzej.swiderski@its.waw.pl
[2]Wojskowa Akademia Techniczna, Wydział Logistyki, Warszawa, ul. Gen. Witolda Urbanowicza 2, 00-908 Warszawa, arkadiusz.jozwiak@wat.edu.pl
[3]Politechnika Warszawska, Wydział Transportu, Warszawa, Plac Politechniki 1, 00-661 Warszawa, rjach@wt.pw.edu.pl

EKSPLOATACYJNE MIARY JAKOŚCI POJAZDÓW W ZASTOSOWANIU DO OCENY USŁUG TRANSPORTOWYCH Z WYKORZYSTANIEM SZTUCZNYCH SIECI NEURONOWYCH

Streszczenie: Eksploatacyjne miary jakości pojazdów są istotnym elementem wykorzystywanym do oceny realizacji usług transportowych. W praktyce mamy do czynienia z wieloma metodami związanymi z eksploatacyjną oceną pojazdów. Scharakteryzowano je w artykule. Metody sztucznej inteligencji, a zwłaszcza sztuczne sieci neuronowe, również mogą być z powodzeniem wykorzystane do tego celu, a zwłaszcza przy podejmowaniu decyzji w procesach oceny jakości maszyn, w tym pojazdów samochodowych. Zastosowanie metod, które pozwalają wspomagać proces decyzyjny na podstawie faktów jest niezmiernie istotne z punktu widzenia wiarygodności i obiektywności oceny. Metody te mogą być również wykorzystane w odniesieniu do eksploatacji pojazdów w zastosowaniu do oceny usług transportowych. W artykule przedstawiono metodę wykorzystania sztucznych sieci neuronowych do eksploatacyjnej oceny pojazdów wykorzystywanych w usługach transportowych towarów. Podstawę weryfikacji metody stanowiły badania eksperymentalne przeprowadzone w przedsiębiorstwie produkującym produkty mleczarskie, współpracującym z firmami transportowymi, dostarczającymi wyroby do produkcji. Uzyskane wyniki potwierdziły z 99-procentowym prawdopodobieństwem wysoką skuteczność proponowanej metody w dokonywaniu oceny usług transportowych z wykorzystaniem eksploatacyjnych miar jakości pojazdów.
Słowa kluczowe: eksploatacja pojazdów, ocena usług transportowych, miary jakości, sztuczne sieci neuronowe

1. Wstęp

Eksploatacyjne miary jakości pojazdów samochodowych służą m.in. wykorzystaniu do oceny realizacji usług transportowych. Istotną grupą problemów dokonania takiej oceny jest dobór odpowiedniej metody. Eksploatacyjna ocena obiektu wymaga zdefiniowania miar (mierników, wskaźników) i określenia ich wartości. Właściwe wartościowanie miar eksploatacyjnych pojazdów jest jednym z kluczowych kryteriów prawidłowego funkcjonowania całego systemu transportowego [9]. Liczbowej oceny sprawności urządzeń dokonuje się w oparciu o wartości, które pochodzą z obserwacji urządzenia podczas eksploatacji [10]. Różnorodność miar eksploatacyjnych zależy oczywiście od typu obiektu (procesu) przy czym, zazwyczaj miary te posiadają różne miana i rzędy skali co powoduje, że są wzajemnie nieporównywalne [6,11]. Porównanie miar opisujących obiekt (proces) jest możliwe dopiero po ich normalizacji. Wśród grup cech obiektów technicznych istotnych dla ich eksploatacyjnej oceny (wyznaczania ich miar i wskaźników) wyróżniono m.in.[8]:

- stan techniczny obiektu, będący miarą możliwości użytkowania obiektu w czasie,

- niezawodność w ujęciu statystycznym,
- jakość, rozumianą jako zdolność obiektu do zaspokojenia określonych potrzeb,
- funkcjonalność opisującą obiekt w sferze kontaktów z człowiekiem,
- efektywność charakteryzującą wydajność obiektu,
- obsługiwalność charakteryzującą podatność obiektu na wykonywanie czynności obsługowych,
- diagnozowalność charakteryzującą podatność obiektu na pozyskiwanie informacji o stanie technicznym.

Wyznaczanie miar wyżej wymienionych grup cech obiektów wymaga stosowania modeli matematycznych. Do najczęściej wykorzystywanych modeli w tym zakresie zalicza się modele niezawodnościowe [8, 14], modele efektywności eksploatacyjnej OEE (Overall Equipment Effectiveness) oraz modele organizacyjno - techniczne KPI (Key Performance Indicators) [17].

Model niezawodnościowy pozwala na statystyczne wyznaczanie miar eksploatacyjnych. Podstawą miar niezawodnościowych w tym modelu jest funkcja niezawodności określana jako prawdopodobieństwo poprawnej pracy obiektu w założonym czasie [14]. W praktyce modele niezawodnościowe umożliwiają określenie wskaźników odnoszących się do obiektów eksploatacji w ujęciu technicznym oraz techniczno - organizacyjnym.

Modele efektywności eksploatacyjnej skupiają miary eksploatowania za pomocą dostępności obiektu, efektywności i jakości jego działania. Model organizacyjno - techniczny KPI (Key Performance Indicators) obejmuje zbiór kluczowych miar wydajności i efektywności. Miary te określone zostały w normie EN 15341:2007 (Maintenance - Maintenance Key Preformance Indicators). Norma ta zawiera 72 wskaźniki wraz ze szczegółową interpretacją elementów, które się na nie składają [15].

Określenie wartości miar cech eksploatacyjnych obiektów umożliwia w dalszej kolejności dokonanie oceny zmian tych wartości w określonym czasie. Dokonywanie oceny obiektów czy procesów wiąże się z podjęciem decyzji. Decyzje zwykle mają prowadzić do zaspokojenia całego zbioru potrzeb decydenta co powoduje konieczność porównywania możliwych rozwiązań, wariantów pod względem wielu kryteriów charakteryzujących dany obiekt czy proces. Stąd podejmowanie złożonych decyzji wymaga zastosowania metod wielokryterialnej analizy (wielokryterialnego podejmowania decyzji, ang. MCDM Multi Criteria Decision Making). Metody te odgrywają ważną rolę m.in. w diagnozie istniejących obiektów czy rozwiązań organizacyjnych [22]. Z uwagi na fakt, iż cechy obiektów czy systemów wyrażane są zazwyczaj w rożnych jednostkach miary, ich stany bezwzględnie nie mogą być ze sobą bezpośrednio porównywane. Dopiero podział zestawu cech charakteryzujących dany obiekt (system), ze względu na pożądane tendencje kształtowania się ich wartości, umożliwia ujednolicenie kryteriów cząstkowych i porównywanie tych cech. Metody wielokryterialnego wspomagania decyzji można podzielić na wywodzące się z teorii użyteczności (m.in.: UTA, UTASTAR, AHP, ANP, SMART) oraz metody oparte na relacji przewyższania (m.in.: ELECTRE, PROMETHEE, ORESTE, REGIME), które wskazują, że ze względu na określone kryterium jedno rozwiązanie jest „co najmniej tak dobre" jak drugie rozwiązanie. Metoda UTA (fr. UTilités Additives) jest oparta na zasadzie agregacji/podziału. Wykorzystuje techniki programowania liniowego w celu optymalnego określenia addytywnych funkcji wartości/użyteczności tak, aby funkcje te były jak najbardziej spójne z preferencjami decydenta [24]. Rozwinięciem metody UTA jest metoda UTASTAR. Zastosowano w niej dodatkowo dwie funkcje błędu oznaczające naruszenie dolnego i górnego krańca funkcji użyteczności grupy alternatyw przez k-ty wariant decyzyjny [24].

Metoda AHP (ang. Analytic Hierarchy Process) jest ogólnym hierarchicznym podejściem do podejmowania wielokryterialnych decyzji, które pozwala łączyć kryteria kwantyfikowane z niekwantyfikowanymi oraz obiektywnie mierzalne z subiektywnymi [11, 17, 18]. Modelowanie za pomocą hierarchicznej analizy problemu AHP jest przydatne szczególnie wtedy, gdy nie jest znana zależność funkcyjna między elementami problemu decyzyjnego,

opisanego w postaci hierarchii czynników, natomiast jest możliwy do oszacowania efekt występowania danych własności i ich efektu praktycznego. Rozszerzeniem metody AHP jest metoda ANP (ang. Analytic Network Process) [1]. Można ją zastosować do rozwiązywania bardziej złożonych problemów decyzyjnych. Konstruowany jest system komponentów istotnych dla rozpatrywanego problemu decyzyjnego w postaci sieci. Uwzględniane są tutaj nie tylko zależności pomiędzy grupami elementów czy wewnątrz nich, lecz także sprzężenia zwrotne.

Metoda SMART (ang. Simple Multi - Attribute Rating Technique) jest metodą wielu etapów, w których identyfikowani są decydenci, możliwości działania, atrybuty istotne dla danego problemu decyzyjnego, wartości i wagi poszczególnych atrybutów, podejmowana jest decyzja oraz wykonywana jest analiza jej wrażliwości [4].

Metody ELECTRE (fr. ELimination Et Choix Traduisant la REalité) [3] mają bardzo szerokie zastosowanie w różnorodnych problemach decyzyjnych. Obejmują one grupę metod (m.in. ELECTRE I, Iv, IS, II, III, IV, TRI) przystosowanych do rozwiazywania różnych problemów wielokryterialnego wspomagania decyzji. Wybór konkretnej metody zależy z jednej strony od rodzaju problemu z jakim mamy do czynienia, z drugiej zaś od rodzaju danych jakimi dysponujemy. W metodach ELECTRE przyjmuje się aksjomat ograniczonej porównywalności wariantów, którego wyrazem jest uznanie czterech podstawowych relacji: I – równoważności, P – silnej preferencji, Q – słabej preferencja oraz R – nieporównywalności. Podstawową zasadą wykorzystywaną w metodach ELECTRE jest porównywanie każdego wariantu z wszystkimi pozostałymi. W ten sposób sprawdza się, czy można uznać dany wariant za mający przewagę nad każdym z pozostałych.

Metody PROMETHEE (ang. Preference Ranking Organization METHod for Enrichment Evaluations) stanowią grupę metod podobnie jak ELECTRE [2]. W metodach PROMETHEE wykorzystuje się informacje na temat stopnia preferencji danego wariantu w stosunku do pozostałych wariantów oraz informacje na temat stopnia, w jakim pozostałe warianty są bardziej preferowane do danego wariantu.

Metoda ORESTE została opracowana dla sytuacji, gdzie alternatywy są uporządkowane według każdego kryterium i same kryteria są uszeregowane według ich znaczenia [12]. W metodzie tej stosuje się niezależne rankingi dla kryteriów oraz dla alternatyw względem każdego z kryteriów.

Metoda REGIME [12] oparta jest na analizie przewyższania i może być postrzegana jako porządkowe uogólnienie metod porównania w parach takich jak analiza zgodności. Podstawą REGIME są współczynniki zgodności Cil zdefiniowane jako suma wag dla kryteriów, dla których alternatywa ai jest przynajmniej tak dobra jak al. Celem tej metody jest określenie znaku różnicy Cil – Cli. Jeżeli znak tej różnicy jest dodatni to alternatywa ai jest preferowana nad alternatywą al i odwrotnie (gdy znak jest ujemny).

Dobór metody jest bardzo istotny z punktu widzenia informacji wyjściowej po dokonaniu oceny. Zależy również od charakteru posiadanych informacji wejściowych, ich ilości oraz wspólnych zależności, jeśli takie występują oraz informacji (celu), jaką chce się uzyskać na wyjściu. Dlatego wykorzystywane do oceny wskaźniki i mierniki poszczególnych elementów eksploatacyjnej oceny pojazdów nie wyczerpują tematu, gdyż przedstawiają jedną zmienną. Nie odzwierciedlają wzajemnych relacji poszczególnych zmiennych i siły (wielkości) tej relacji.

Do eksploatacyjnej oceny pojazdów można również wykorzystać metody, zaliczane do sztucznej inteligencji, w tym sztuczne sieci neuronowe.

Celem artykułu jest prezentacja możliwości wykorzystania sztucznych sieci neuronowych do:

- wspomagania decyzji, związanych z eksploatacją pojazdów, wykorzystywanych w usługach transportowych, związanych z dostawami wyrobów do produkcji,

- prognozowania jakości i efektywności eksploatacyjnej pojazdów samochodowych w systemie usług transportowych.

Zastosowano takie metody badawcze, jak: analiza (wykorzystana do rozpoznania obszaru sztucznej inteligencji), modelowanie opisowe (posłużyło do sformułowania i opisania zebranych informacji), modelowanie matematyczne, wykorzystujące sztuczne sieci neuronowe (do eksploatacyjnej oceny pojazdów samochodowych).

2. Aspekty eksploatacyjnej oceny usług transportowych

Usługi transportowe są niezmiennie istotnym elementem gospodarki i życia społecznego, umożliwiającym skuteczne ich funkcjonowanie. Rozwój społeczno - gospodarczy generuje potrzebę przemieszczania osób i/lub ładunków. Brak spójności między działalnością transportową i wytwórczą zdecydowanie osłabia możliwości rozwojowe. Dodatkowo, duża konkurencja w tym segmencie spowodowała, że najniższa cena przestała być gwarantem przewagi rynkowej. Rozważania te, to jedne z licznych aspektów powodujących zainteresowanie oceną usług transportowych w aspekcie eksploatacji pojazdów. Problem jakości w perspektywie dynamicznych zmian na rynku, staje się szczególnie ważny z takich względów, jak: nieustanny wzrost oczekiwań klientów, minimalizacja czasu trwania usługi, gwarancja najwyższej efektywności usługi, czy też bezpieczeństwo pojazdu i przewożonego towaru. Jednym z istotniejszych wymiarów oceny usługi transportowej jest ocena z punktu widzenia eksploatacji pojazdu. Ocena ta, to problem złożony ze względu na liczne kryteria, opisane atrybutami niemierzalnymi lub trudno mierzalnymi. Z reguły kryteria mają charakter heterogeniczny, co dodatkowo komplikuje wiarygodną ocenę. Problem ten rozwiązuje zastosowanie wielokryterialnych metod podejmowania decyzji w oparciu np. o metody heurystyczne czy teorie zbiorów rozmytych. Niestety, metody te ze względu na swoją konstrukcję matematyczną są trudne do implementacji. Z tego powodu poszukuje się prostych, praktycznych narzędzi, dających utylitarne korzyści. W sytuacji, kiedy występuje pełna znajomość reguł i mała złożoność problemu, zastosowanie znajdują dokładne algorytmy (np. modele liniowe). W przypadku częściowej lub całkowitej nieznajomości reguł i dużej złożoności problemu, zastosowanie znajdują sieci neuronowe. Ocena eksploatacyjna, ze względu na swoją złożoność i wieloaspektowość, należy do obszaru zastosowań sztucznej inteligencji.

Istnieje wiele kryteriów klasyfikacji usług transportowych. Do istotnych, które wpływają na obszary oceny, można zaliczyć: przedmiot przewozu (np. rynek przewozów pasażerskich, rynek przewozów towarowych), wykorzystywany rodzaj transportu, obszar działania (np. rynek lokalny, krajowy, międzynarodowy), siłę ekonomiczną podmiotów (rynek przewoźnika lub rynek użytkownika), itp.

W artykule podjęto próbę dokonania eksploatacyjnej oceny, której przedmiotem badań są pojazdy samochodowe.

Podstawowe cechy usług transportowych, to m.in.:

- złożoność - usługa transportowa składa się z bardzo dużej liczby elementów i relacji między tymi elementami,
- probabilizm - nie można przewidzieć wszystkich stanów i zdarzeń,
- dynamiczność - na bieżąco ingeruje się, zarówno w czasie, jak i w przestrzeni w proces realizacji usługi transportowej.

Do podstawowych uwarunkowań funkcjonowania usług transportowych można zaliczyć:

- aspekty ekonomiczno - prawne - np. system finansowy, akty prawne regulujące transport,

- aspekty techniczne - np. pojazdy, infrastruktura, urządzenia transportowe i przeładunkowe, wraz z wieloma aspektami eksploatacyjnymi,
- aspekty organizacyjne - np. zasady współpracy między przewoźnikiem a klientem, czas pracy przewoźnika.

Aspekty jakościowe realizacji usług transportowych stanowią odrębną grupę obszarów badawczych. Wyróżnić można trzy kategorie jakości [23]:

- postulowaną przez użytkowników, która określa ich żądania i życzenia dotyczące sposobu realizacji usługi transportowej,
- oferowaną przez przewoźników, czyli oferowaną podaż usług transportowych możliwych do realizacji przy aktualnym zasobie wiedzy, techniki i organizacji,
- realizowaną przez usługodawców.

Zagadnienia oceny usług transportowych rozpatruje się w różnych ujęciach. Najczęściej analizy tej dokonuje się ze względu na czas dostawy, bezpieczeństwo i niezawodność realizacji usługi oraz bezpieczeństwo i niezawodność samych pojazdów [5]. Z kolei Neo i inni dokonali analizy jakości usług świadczonych przez operatorów logistycznych i jako podstawowe wskaźniki oceny podali dokładność informacji, dokładność realizacji procesów kompletacji i terminową realizację dostaw [13].

Biorąc powyższe pod uwagę, oceny usługi transportowej można dokonać pod wieloma innymi aspektami, tj.: poniesionych kosztów, ryzyka, zasobów: ludzkich, informacyjnych czy eksploatowanych pojazdów. W zależności od wspomnianych wcześniej uwarunkowań i charakteru wykonywanej usługi aspekty te są różnie interpretowane. Chociaż w przypadku np. szacowania ryzyka, zaleca się przeprowadzenie tego procesu w określonej kolejności: określenie zakresu, identyfikacja zagrożeń i wstępne wyznaczenie konsekwencji, oszacowanie ryzyka, weryfikacja, dokumentowanie i uaktualnianie analizy [16].

Ocenę usługi transportowej można przedstawić jako funkcję postaci:

$$R_n(t)=f(w_{n,1}(t),\ w_{n,2}(t),\ldots,\ w_{n,k}(t)) \qquad (1)$$

gdzie: $R_n(t)$ - ocena n-tej usługi transportowej w czasie t,
$w_{n,k}(t)$ - ocena k-tego wymagania n-tej usługi transportowej w czasie t.

Aspekty przedstawione wyżej mogą być opisane różnymi miarami w zależności, z którego punktu widzenia dokonuje się oceny. Natomiast istotne są niezawodne, eksploatowane w trakcie realizacji usługi transportowej pojazdy. Ich jakość zależy w głównej mierze od właściwej eksploatacji, która opisywana jest przez niezawodność i gotowość. Z kolei na gotowość pojazdu składają się takie elementy, jak: nieuszkadzalność, obsługiwalność i zapewnienie środków obsługi [15]. Dlatego, na potrzeby niniejszej pracy eksploatacyjną ocenę n-tej usługi transportowej zdefiniowano jako funkcję:

$$E_n(t) = f\,(w_{n,u}(t),\ w_{n,o}(t),\ w_{n,w}(t),\ w_{n,st}(t)) \qquad (2)$$

gdzie: $E_n(t)$ - eksploatacyjna ocena *n-tej* usługi transportowej w czasie t,
$w_{n,u}(t)$ - ocena wymagania nieuszkadzalności pojazdu wykonującego n-tą usługę transportową w czasie t,
$w_{n,o}(t)$ - ocena wymagania obsługiwalności pojazdu wykonującego n-tą usługę transportową w czasie t,
$w_{n,w}(t)$ - ocena wymagania dotycząca wieku pojazdu wykonującego n-tą usługę transportową w czasie t,
$w_{n,st}(t)$ - ocena wymagania stanu technicznego pojazdu wykonującego n-tą usługę transportową w czasie t.

Do dokonania oceny eksploatacyjnej usług transportowych, ze względu na swoją wielopłaszczyznowość, potrzebne są narzędzia, które odnajdą zależności między zbiorami zmiennych przy dużej złożoności problemu. Narzędziem tym mogą być sztuczne sieci neuronowe. Zatem w dalszej części artykułu przedstawiono modelowanie, z wykorzystaniem tego narzędzia, oceny eksploatacyjnej pojazdów w zastosowaniu do oceny usług transportowych oraz wyniki badań własnych.

3. Modelowanie neuronowe

Badania dotyczące możliwości i sposobu wykorzystania sztucznych sieci neuronowych do eksploatacyjnej oceny usług transportowych przeprowadzono na bazie usług świadczonych przez zewnętrznych przewoźników na rzecz przedsiębiorstwa produkującego i dostarczającego na rynek wyroby mleczarskie. Ocenie poddano usługi transportowe realizowane na rynku krajowym o zasięgu lokalnym z wykorzystaniem pojazdów samochodowych.

W celu zastosowania sieci neuronowych do dokonania eksploatacyjnej oceny usług transportowych określono zbiór sygnałów wejściowych, zarówno ilościowych (wyrażonych w postaci liczby), jak i jakościowych (wyrażonych w postaci opisu, słowa). Miary jakości pojazdów, opisujące wymagania eksploatacyjnej oceny usług transportowych przedstawiono w tabeli 1.

Tabela. 1. Opis wymagań eksploatacyjnej oceny usługi transportowej E_n

Symbol wymagania oceny	**Opis wymagania eksploatacyjnej oceny usługi transportowej E_n**
$w_{n,u}(t)$	nieuszkadzalność pojazdu - liczba uszkodzeń występujących w jednostce czasu (np. w tygodniu, miesiącu, itp.),
$w_{n,o}(t)$	obsługiwalność pojazdu – liczba godzin pojazdu w stanie obsługi,
$w_{n,w}(t)$	wiek pojazdu - przedkłada się na pozostałe charakterystyki niezawodnościowe, dlatego wymaganie to ujęto w miarach jakości pojazdu,
$w_{n,st}(t)$	stan techniczny pojazdu - charakterystyka ta wynika m.in. z pozostałych wskaźników oceny niezawodności i oceniania jest przez eksperta organoleptycznie, przyjęto oceny: bardzo dobry stan techniczny pojazdu (bdb), zadawalający (zad), niezadawalający (nzad).

Źródło: opracowanie własne.

Dla potrzeb oceny jakości usługi transportowej $E_n(t)$ wg wymagań $w_{n,u}(t)$, $w_{n,o}(t)$, $w_{n,w}(t)$ i $w_{n,st}(t)$, dokonano ich parametryzacji zgodnie z ustaleniami podanymi w tabelach 2 - 5.

Tabela. 2. Sparametryzowana ocena wymagania $w_{n,u}$ - nieuszkadzalność pojazdu

Lp.	**Nieuszkadzalność pojazdu [liczba/miesiąc]**	**Sparametryzowana ocena jakości**	**Opisowa ocena jakości**
1	0	1	wysoki poziom jakości
2	1	0,75	bardzo dobry poziom jakości
3	2	0,5	dobry poziom jakości
4	3	0,25	niski poziom jakości
5	4	0	nieakceptowalny poziom jakości

Źródło: opracowanie własne.

Tabela. 3. Sparametryzowana ocena wymagania $w_{n,o}$ - obsługiwalność pojazdu

Lp.	Obsługiwalność pojazdu [liczba godzin]	Sparametryzowana ocena jakości	Opisowa ocena jakości
1	0	1	wysoki poziom jakości
2	0-1	0,75	bardzo dobry poziom jakości
3	2-5	0,5	dobry poziom jakości
4	6-10	0,25	niski poziom jakości
5	>10	0	nieakceptowalny poziom jakości

Źródło: opracowanie własne.

Tabela. 4. Sparametryzowana ocena wymagania $w_{n,w}$ – wiek pojazdu

Lp.	Wiek pojazdu [w latach]	Sparametryzowana ocena jakości	Opisowa ocena jakości
1	0-5	1	wysoki poziom jakości
2	6-12	0,5	dobry poziom jakości
3	>12	0	nieakceptowalny poziom jakości

Źródło: opracowanie własne.

Tabela. 5. Sparametryzowana ocena wymagania $w_{n,st}$ – stan techniczny pojazdu

Lp.	Stan techniczny pojazdu	Sparametryzowana ocena jakości	Opisowa ocena jakości
1	bardzo dobry (bdb)	1	wysoki poziom jakości
3	zadawalający (zad)	0,5	dobry poziom jakości
5	nie zadawalający (nie)	0	nieakceptowalny poziom jakości

Źródło: opracowanie własne.

Do dokonywania oceny poszczególnych wymagań, eksperci przypisali wagi charakterystykom w skali (0-10), gdzie 0 oznacza mało istotne, 10 oznacza bardzo istotne. Wymaganiom przypisano następujące wartości wag:

- $w_{n,u} – 8$,
- $w_{n,o} – 3$,
- $w_{n,w} – 5$,
- $w_{n,st} – 5$.

Przy tak zdefiniowanych wymaganiach oceny eksploatacyjnej, użytkownicy pojazdów i eksperci dokonywali oceny poszczególnych wymagań. Zebrano dane z eksploatacji (z realizacji 812 usług transportowych zrealizowanych w ciągu 5 ostatnich lat dla potrzeb dostaw wyrobów do produkcji). Dane te stanowiły punk wyjścia do rozpoczęcia badań (tabela 6). Na podstawie średniej ważonej poszczególnych wymagań dokonano eksploatacyjnej oceny jakości usług transportowych (pozytywna lub negatywna). Poziom jakości uznany jako zadawalający przyjęto na poziomie 0,6 i więcej.

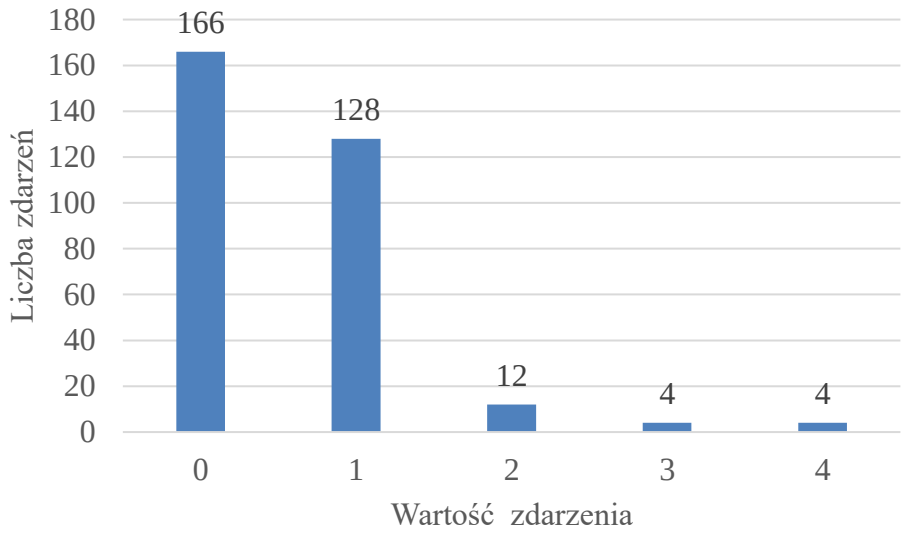
Tabela. 6. Przykładowe dane do uczenia sieci neuronowej

Lp.	Nieuszkadzalność pojazdu	Obsługiwalność pojazdu	Wiek pojazdu	Stan techniczny pojazdu	Ocena (średnia ważona)	Ocena
	liczba/miesiąc	liczba godzin/miesiąc	lata	bdb/zad/nzad		
1.	0	0	1	bdb	**1,00**	**Pozytywna**
2.	1	1	3	bdb	**0,87**	**Pozytywna**
3.	0	0	6	bdb	**0,88**	**Pozytywna**

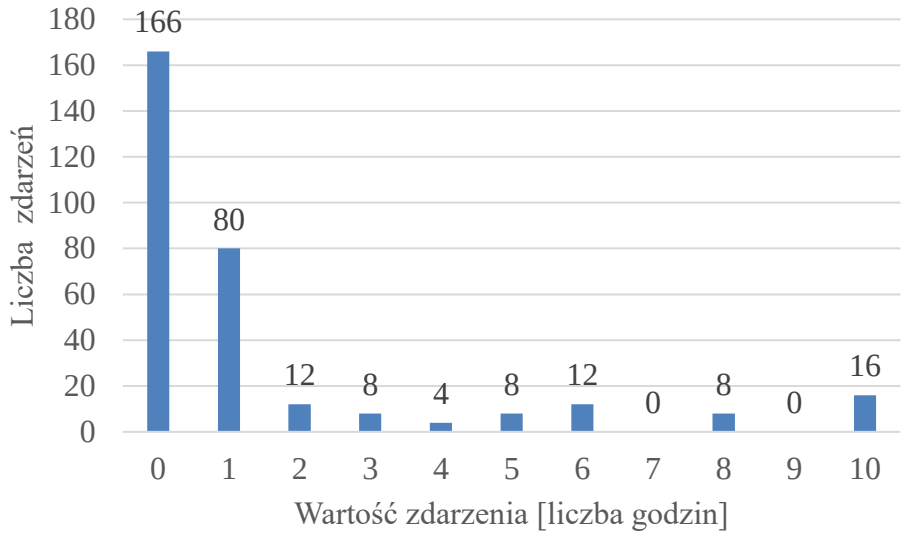
Lp.	Nieuszkadzalność pojazdu	Obsługiwalność pojazdu	Wiek pojazdu	Stan techniczny pojazdu	Ocena (średnia ważona)	Ocena
	liczba/miesiąc	liczba godzin/miesiąc	lata	bdb/zad/nzad		
4.	1	1	6	bdb	**0,75**	**Pozytywna**
5.	1	1	7	zad	**0,63**	**Pozytywna**
6.	0	0	7	zad	**0,76**	**Pozytywna**
7.	1	6	8	zad	**0,56**	**Negatywna**
8.	0	0	8	zad	**0,76**	**Pozytywna**
9.	0	0	8	bdb	**0,88**	**Pozytywna**
10.	1	6	11	nzad	**0,44**	**Negatywna**
11.	0	0	11	nazd	**0,64**	**Pozytywna**
12.	2	8	12	zad	**0,46**	**Negatywna**
13.	0	0	13	zad	**0,64**	**Pozytywna**
14.	1	2	13	nazd	**0,36**	**Negatywna**

Źródło: opracowanie własne.

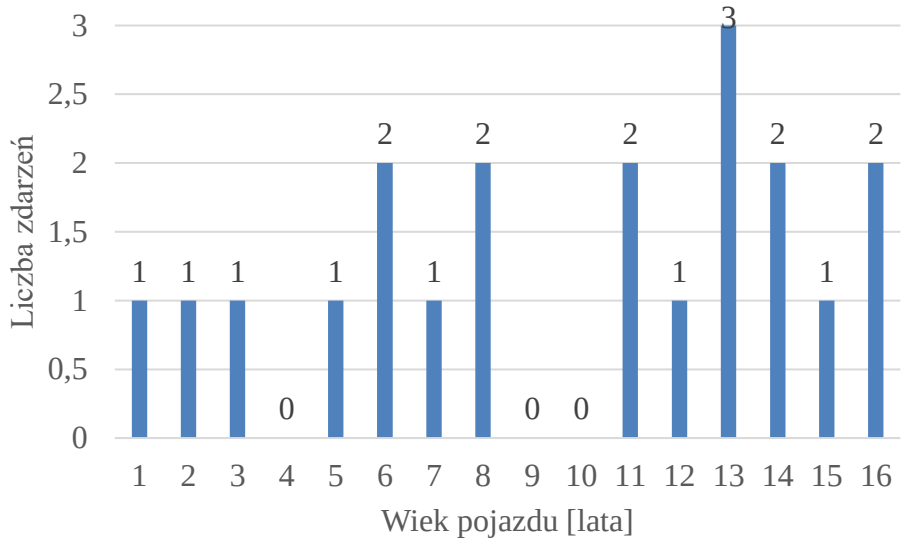
Poniżej przedstawiono strukturę otrzymywanych wyników badań, wg poszczególnych wymagań (rys. 1-4).



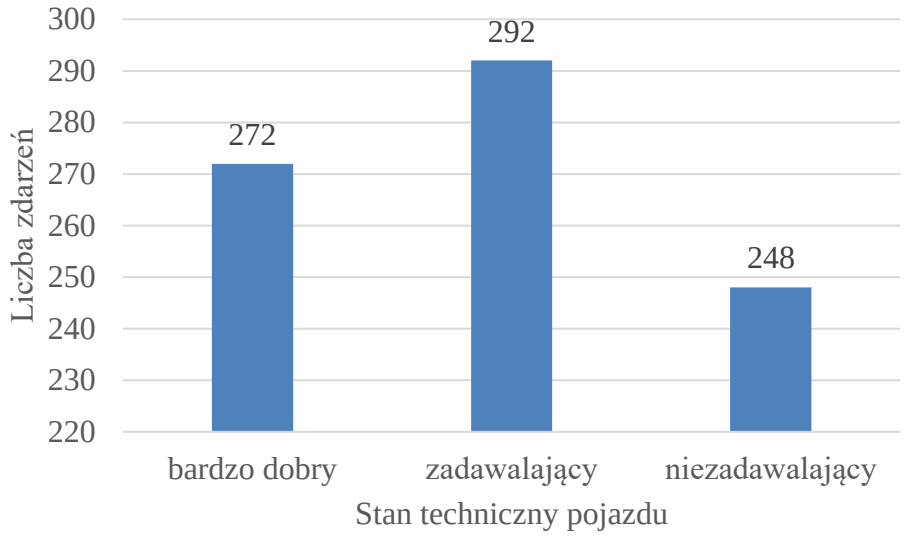
Rys. 1. Wyniki badań wymagania $w_{n,u}$ – nieuszkadzalność pojazdu
Źródło: opracowanie własne.



Rys. 2. Wyniki badań wymagania $w_{n,o}$ – obsługiwalność pojazdu
Źródło: opracowanie własne.



Rys. 3. Wyniki badań wymagania $w_{n,w}$ – wiek pojazdu
Źródło: opracowanie własne.



Rys. 4. Wyniki badań wymagania $w_{n,st}$ – stan techniczny pojazdu
Źródło: opracowanie własne.

Na podstawie powyższych danych otrzymano 540 ocen pozytywnych i 272 ocen negatywnych.

Spośród wielu rodzajów sieci neuronowych oraz wielu algorytmów ich uczenia w dalszych badaniach wykorzystano perceptron wielowarstwowy (Multilayer Perceptron) i algorytmy uczące: metoda najszybszego spadku, metoda gradientów sprzężonych; metodę BFGS (*Broyden – Fletcher – Goldfarb - Shanno)*. Wykorzystana sieć neuronowa należy do grup:

- tzw. sieci nadzorowanych, gdzie proces uczenia odbywa się pod nadzorem nauczyciela (wśród sygnałów wychodzących jest sygnał wzorcowy),
- sieci jednokierunkowych, gdzie przepływ sygnałów (informacji) odbywa się w jedną stronę (od wejścia do wyjścia sieci neuronowej).

Wykorzystując program komputerowy Statistica 12 dokonano oceny usług transportowych z wykorzystaniem zdefiniowanych wcześniej eksploatacyjnych miar jakości pojazdów.

Określono zatem następujące sygnały:

- wejściowe ilościowe: $w_{n,u}(t)$, $w_{n,o}(t)$, $w_{n,w}(t)$,
- wejściowe jakościowe: $w_{n,st}(t)$,
- wyjściowe ilościowe: $E_n(t)$.

Wraz ze wskazaniem danych wejściowych, zdefiniowano wielkości zbiorów. Określono, że:

- 80 % - danych stanowić będzie 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.

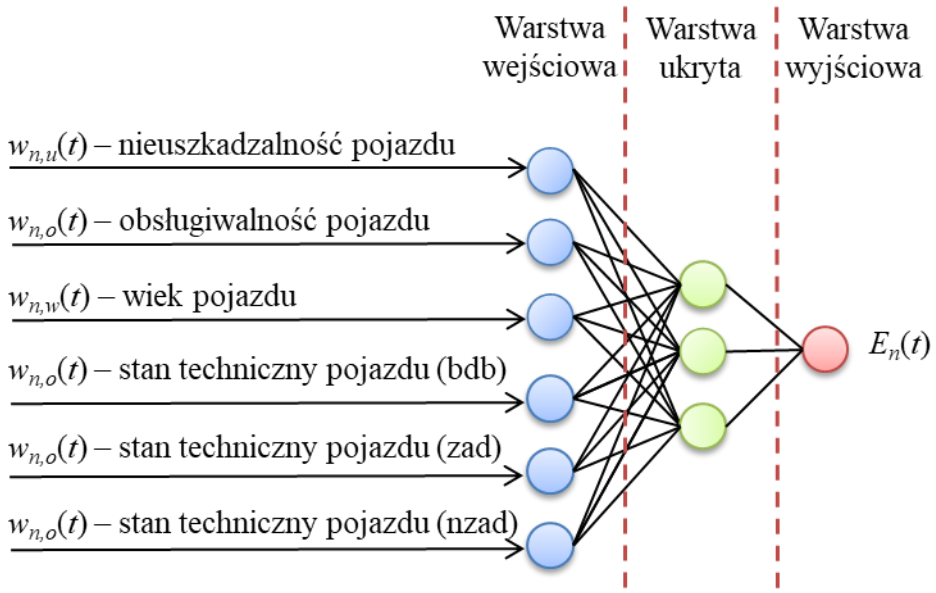
Następnie sprecyzowano podstawowe parametry sieci, tj.:

- typ sieci (perceptron wielowarstwowy - MLP),
- minimalną liczbę neuronów ukrytych,
- maksymalną liczbę neuronów ukrytych,
- liczbę sieci uczących,
- liczbę sieci zachowanych,
- funkcję aktywacji neuronów ukrytych,
- funkcję aktywacji neuronów wyjściowych,

- wartości redukcji wag dla warstwy ukrytej i warstwy wyjściowej.

Po zdefiniowaniu danych i parametrów sieci przeprowadzono proces uczenia sieci neuronowej wykorzystując zgromadzone dane. Przykładowe wyniki tego procesu przedstawiono w tabeli 7.

Przy tak określonych wymaganiach i przeprowadzeniu procesu uczenia, struktura najlepszej sieci przyjęła postać MLP 6-3-1, co oznacza 6 neuronów w warstwie wejściowej, 3 neurony w warstwie ukrytej i 1 neuron w warstwie wyjściowej (rys. 5).



Rys. 5. Struktura badanej sieci MLP 6-3-1
Źródło: opracowanie własne.

Jakość uczenia sieci MLP 6-3-1 została oszacowana na poziomie 99,6% prawdopodobieństwa wskazania poprawnej oceny (odpowiedzi), jakość testowania na poziomie 99,7%, co oznacza, że prawie wszystkie próby w tym zbiorze zostały prawidłowo przyporządkowane i jakość walidacji określono na poziomie 99,4%. Najlepszym algorytmem uczenia okazał się algorytm BFGS 148 (liczba 148 oznacza, liczbę epok, jakich sieć potrzebowała do przeprowadzenia procesu uczenia i odnalezienia najlepszej sieci, w której błąd uczenia był najmniejszy).

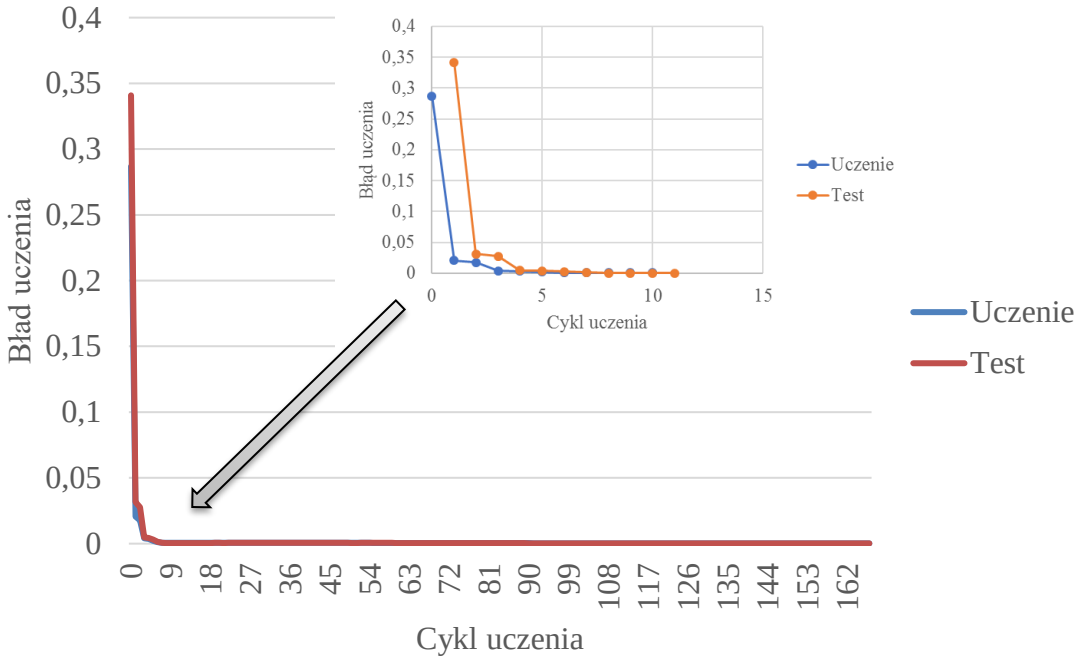
Tabela. 7. Przykładowe wyniki procesu uczenia sieci neuronowej

ID	Nazwa sieci	Jakość uczenia	Jakość testowania	Jakość walidacji	Algorytm uczenia	Funkcja błędu	Aktywacja ukryta	Aktywacja wyjściowa
1	MLP 6-5-1	0,988	0,992	0,989	BFGS 55	SOS	Logistyczna	Logistyczna
2	MLP 6-4-1	0,988	0,992	0,989	BFGS 98	SOS	Logistyczna	Tanh
3	MLP 6-6-1	0,984	0,990	0,991	BFGS 10	SOS	Liniowa	Liniowa
4	MLP 6-7-1	0,984	0,991	0,990	BFGS 8	SOS	Liniowa	Liniowa
5	MLP 6-8-1	0,983	0,991	0,991	BFGS 10	SOS	Liniowa	Liniowa
6	MLP 6-4-1	0,983	0,992	0,991	Najszybszego spadku 30	SOS	Tanh	Liniowa
7	MLP 6-4-1	0,987	0,993	0,989	BFGS 47	SOS	Logistyczna	Wykładnicza
8	MLP 6-8-1	0,987	0,992	0,987	BFGS 48	SOS	Logistyczna	Logistyczna
9	MLP 6-7-1	0,990	0,993	0,989	BFGS 63	SOS	Logistyczna	Liniowa
10	MLP 6-3-1	0,996	0,997	0,994	BFGS 148	SOS	Tanh	Liniowa
11	MLP 6-8-1	0,983	0,991	0,991	BFGS 9	SOS	Liniowa	Liniowa
12	MLP 6-4-1	0,985	0,992	0,988	Gradienty sprzężone 27	SOS	Tanh	Liniowa
13	MLP 6-6-1	0,977	0,984	0,990	BFGS 12	SOS	Logistyczna	Sinus
14	MLP 6-4-1	0,981	0,984	0,990	BFGS 23	SOS	Sinus	Sinus
15	MLP 6-6-1	0,981	0,988	0,990	BFGS 23	SOS	Sinus	Sinus
16	MLP 6-10-1	0,996	0,995	0,997	BFGS 87	SOS	Tanh	Liniowa
17	MLP 6-5-1	0,991	0,995	0,990	BFGS 86	SOS	Tanh	Tanh
18	MLP 6-8-1	0,985	0,992	0,988	BFGS 16	SOS	Wykładnicza	Logistyczna
19	MLP 6-14-1	0,991	0,993	0,990	BFGS 47	SOS	Tanh	Wykładnicza
20	MLP 6-1-1	0,976	0,983	0,990	BFGS 16	SOS	Liniowa	Sinus
21	MLP 6-1-1	0,985	0,992	0,987	BFGS 39	SOS	Wykładnicza	Wykładnicza

Źródło: opracowanie własne.

4. Weryfikacja wybranej sieci neuronowej MLP 6-3-1

O pozytywnym wyniku uczenia sieci neuronowej świadczy m.in. wykres uczenia (rys. 6), z którego wynika, że najlepszą strukturę sieci odnaleziono w 148 epoce, w której udział błędnych odpowiedzi wynosi poniżej 1 %, a błąd został oszacowany na poziomie 0,0002.



Rys. 6. Wykres uczenia sieci neuronowej MLP 6-3-1
Źródło: opracowanie własne.

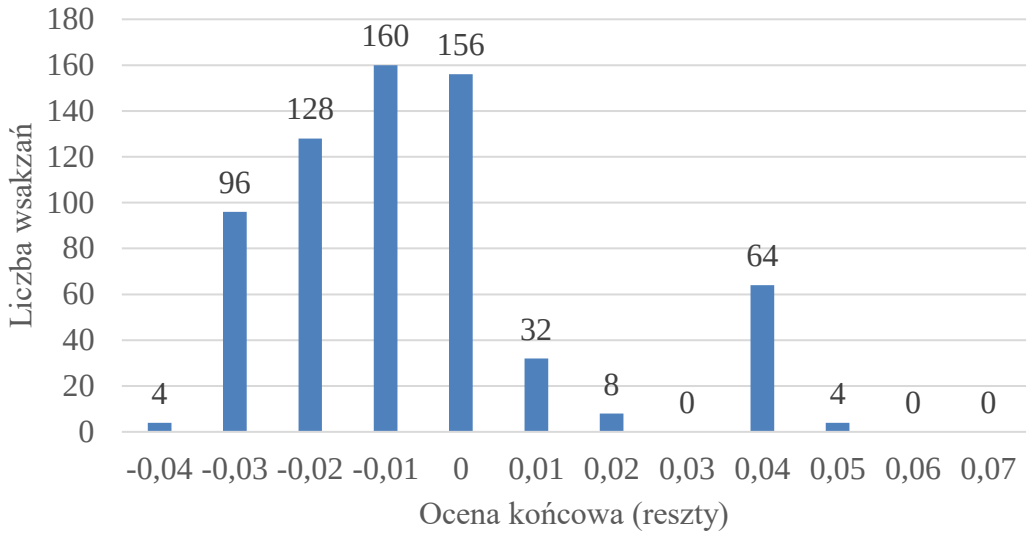
Macierz pomyłek przedstawiono w tabeli 8. Wskazuje ona dokładnie, ile przypadków danej oceny zostało zakwalifikowanych przez sieć do oceny pozytywnej (powyżej progu 0,6) lub negatywnej (poniżej progu 0,6). Z tabeli 8 wynika, iż na 432 oceny pozytywne, sieć prawidłowo przyporządkowała 408 wskazań, natomiast prawidłowo wskazała wszystkie oceny negatywne.

Tabela. 8. Macierz pomyłek sieci neuronowej MLP 6-3-1

Lp.		Decyzja - negatywna	Decyzja - pozytywna	Decyzja - wszystkie
1	Razem	220	432	652
2	Poprawne	220	408	628
3	Niepoprawne	0	24	24
4	Poprawne (%)	100%	94%	96%
5	Niepoprawne (%)	0%	6%	4%

Źródło: opracowanie własne.

Kolejną istotną cechą badanej sieci neuronowej jest rozkład reszt przedstawiony na rys. 7, czyli różnic między zmienną wyjściową i jej predykcją.



Rys. 7. Rozkład reszt sieci neuronowej MLP 6-3-1

Źródło: opracowanie własne.

Z histogramu można odczytać, iż reszty mają rozkład normalny wokół zera z naciskiem na wartości ujemne. Zdecydowana większość ocen została dokonana z błędem na poziomie -0,03-0.

Ostatnim etapem weryfikacji sieci neuronowej są przewidywania dla nowych danych wejściowych. Prób, które do tej pory nie pojawiły się w żadnym zbiorze. Aby otrzymać nowe eksploatacyjne oceny jakości usługi transportowej, uzupełniono wartości wszystkich sygnałów wejściowych do sieci neuronowej, na podstawie których został wygenerowany sygnał wyjściowy, czyli eksploatacyjna ocena (ilościowa) usługi transportowej. Po wprowadzeniu danych do sieci, otrzymano wyniki końcowe. Tabela 9 zawiera przewidywania dla nowych danych na podstawie sieci neuronowej MLP 6-3-1.

Tabela. 9. Przewidywania dla nowych danych sieci MLP 6-3-1

Lp.	Nieuszkadzalność pojazdu	Obsługiwalność pojazdu	Wiek pojazdu	Stan techniczny pojazdu	Ocena końcowa
1	0	0	3	bdb	0,99
2	1	5	3	bdb	0,863
3	1	2	5	zad	0,726
4	2	10	7	zad	0,552
5	1	5	10	nzad	0,458

Źródło: opracowanie własne.

Uzyskane wyniki wskazują na możliwość wykorzystania sieci neuronowych, w tym przypadku jednowarstwowej sieci wielowarstwowej do dokonywania eksploatacyjnej oceny jakości usług transportowych. Zarówno liczba, jak i rodzaj danych (ilościowe lub jakościowe) nie wpływają na osiągnięcie wysokich wyników na poziomie 98-99% skuteczności. Na podstawie zgromadzonych danych z przeszłości, sieć neuronowa umożliwia podejmowanie decyzji, generowanie oceny z bieżącej lub przyszłej eksploatacji.

5. Podsumowanie

W literaturze przedmiotu badań zdefiniowanych jest wiele różnych metod i modeli (najczęściej wielokryterialnych) do dokonywania oceny eksploatacji pojazdów i systemów technicznych. Scharakteryzowano je we wstępie artykułu. Są one w praktyce powszechnie wykorzystywane. Doświadczenie autorów oraz przeprowadzona analiza stanu wskazuje, że sieci neuronowe nie są jeszcze wykorzystywane powszechnie do eksploatacyjnej oceny pojazdów w zastosowaniu do usług transportowych. Sieci neuronowe są stosowane w eksploatacji, lecz w innym zakresie [7, 19, 20, 21].

Uzyskane wyniki przeprowadzonych badań w przedsiębiorstwie produkującym wyroby mleczarskie oraz w firmach transportowych kooperujących z tym przedsiębiorstwem wskazują na możliwość zastosowania sztucznych sieci neuronowych do dokonywania oceny pojazdów wykorzystywanych przy realizacji usług transportowych, z wykorzystaniem eksploatacyjnych miar jakości. W przypadku negatywnych ocen istnieje możliwość bezzwłocznego podjęcia działań naprawczych. Na tej podstawie można prognozować przyszłe przewidywania dla danych z bieżącej eksploatacji.

Zaproponowane eksploatacyjne miary jakości wynikały z potrzeb badanego przedsiębiorstwa. W praktyce mogą być one dobierane w różny sposób, w zależności od celu prowadzonych analiz.

Sieci neuronowe okazały się więc przydatne jako narzędzie do:

- wspomagania decyzji, związanych z eksploatacją pojazdów, wykorzystywanych w usługach transportowych, związanych z dostawami wyrobów do produkcji,
- prognozowania jakości i efektywności eksploatacyjnej pojazdów samochodowych w systemie usług transportowych.

Zatem cel przeprowadzonych badań został osiągnięty.

Zastosowanie sieci neuronowych w eksploatacji może być szersze. Na przykład do oceny ryzyka nie tylko eksploatacyjnego, ale też związanego z niezawodnością i bezpieczeństwem użytkowania pojazdów (nie tylko samochodowych) i innych maszyn oraz z bezpieczeństwem realizacji samych usług transportowych.

Przykład obliczeniowy zaprezentowany w artykule oddaje zatem istotę zastosowania sztucznych sieci neuronowych do oceny zagadnień związanych z eksploatacją. Jest to jeden z możliwych punktów wyjścia do prowadzenia dalszych badań w perspektywie zastosowania sieci neuronowych w tym obszarze.

LITERATURA:

[1] Adamus W., Gręda A., Wspomaganie decyzji wielokryterialnych w rozwiązywaniu wybranych problemów organizacyjnych i menedżerskich, „Badania operacyjne i decyzje" 2005, Nr 2.

[2] Brans J.P., Mareschal B., Promethee methods, [in:] Figueira J., Greco S., Ehrgott M. (ed.), Multiple Criteria Decision Analysis: State of the Art Surveys, Springer Verlag, Boston, Dordrecht, London, 2005.

[3] Figueira J., Mousseau V., Roy B. ELECTRE methods, [in:] J. Figueira, S. Greco, and M. Ehrgott (eds.), Multiple Criteria Decision Analysis: State of the Art Surveys, Springer Verlag, Boston, Dordrecht, London, 2005.

[4] Goodwin P., Wright G., Analiza decyzji, Wolters Kluwer Polska, Warszawa 2011 Hurley W.J., The analytic hierarchy process: a note on an approach to sensitivity which preserves rank order, „Computers & Operations Research" 2001, No. 28.

[5] Grigoroudis E., Siskos Y., A survey customer satisfaction barometers: Some results from the transportation-communications sector. European Journal of Operational Research 152, 2004, pp. 334.353.
[6] Jacyna-Gołda I, Lewczuk K., The method of estimating dependability of supply chain elements on the base of technical and organizational redundancy of process. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 382–392, http://dx.doi.org/10.17531/ein.2017.3.9.
[7] Karpenko M., Sepehri (2002), Neural network classifiers applied to condition monitoring of a pneumatic process valve actuator. Engineering Applications of Artificial Intelligence, Vol. 15, No. 3, pp. 273-283.
[8] Kaźmierczak J., Eksploatacja systemów technicznych. Gliwice: Wydawnictwo Politechniki Śląskiej, 2000.
[9] Knopik L, Migawa K., Multi-state model of maintenance policy. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (1): 125–130, http://dx.doi.org/10.17531/ein.2018.1.16.
[10] Kornacki A., Sokołowska E., The estimation of smooth operation time until failure with the application of the Akaike Information Criterion (AIC). Eksploatacja i Niezawodnosc - Maintenance and Reliability 2010; 1 (45): 69-76.
[11] Loska A., Exploitation assessment of selected technical objects using taxonomic methods, Eksploatacja i Niezawodność - Maintenance and Reliability 2013, vol 15, No.1.
[12] Martel J.M., Other Outranking Methods, [w:] Multiple Criteria Decision Analysis: State of the Art Surveys, J. Figueira, S. Greco, and M. Ehrgott, editors, Springer Verlag, Boston, Dordrecht, London, 2005.
[13] Neo H.-Y., Xie M., Tsui K._l., Service quality analysis: case study of a 3PL company. International Journal of Logistics Systems and Management 1(1), 2004, p. 64-80.
[14] Niebel W.B. Engineering Maintenance Management. Second edition. New York: Marcel Dekker Inc., 1994.
[15] PN-EN 15341:2007 - Maintenance - Maintenance Key Performance Indicators.
[16] PN-EN 60300-1:2015 Dependability management – Part 1: Guidance for management and application.
[17] Saaty T. Decision making for Leaders The Analytic Hierarchy Process for decisions in a complex world, University of Pittsburgh, RWS Publications, Pittsburgh 2001.
[18] Saaty T. The Analytic Hierarchy Process. RWS Publications. Pittsburgh 1998.
[19] Samanta B. (2004), Gear Fault Detection Using Artificial Neural Network and Support Vector Machines with Genetic Algorithms. Mechanical Systems and Signal Processing, Vol. 18, pp. 625-644.
[20] Samanta B., Al-Balushi K.R. (2003) Artifricial neural network based fault diagnostics of rolling element bearings using time-domain features. Mechanical Systems and Signal Processing, Vol. 17, Issue 2, pp. 317-328.
[21] Samanta B., Al-Balushi K.R., Al-Araimi S.A. (2001), Use of genetic algorithm and artificial naural network form gear condition diagnostics. Proc. of COMADEM, University of Manchester, pp. 449-456.
[22] Szudrowicz M., Layered composite increasing the resistance of patrol and intervention vehicles to the impact of improvised explosive devices (iEd) from below. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20 (1): 9–15, http://dx.doi.org/10.17531/ein.2018.1.2.
[23] Świderski A, *Modelowanie oceny jakości usług transportowych,* Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, 2011.
[24] Yannis Siskos, Evangelos Grigoroudis, Nikolaos F Matsatsinis, Multiple criteria decision analysis: State of the art surveys, Springer New York, 2005.

Ewelina Kilian – Błażejewska

Instytut Inżynierii Wody i Ścieków
Wydział Inżynierii Środowiska i Energetyki
Politechnika Śląska
Akademicka 2A, 44 – 100 Gliwice, Poland
e – mail: eewelina.kilian@gmail.com

Wpływ wybranych czynników na intensywność uszkodzeń przewodów podsystemu dystrybucji wody, zlokalizowanych na terenie oddziaływania wstrząsów pochodzenia górniczego

The influence of selected factors on the failure rate of the pipelines located on the area of the impact of mining tremors

Słowa kluczowe: *podsystem dystrybucji wody, wstrząsy górnicze, intensywność uszkodzeń przewodów sieci wodociągowej*

Keywords*: Water distribution system, mining tremors, failure rate of water pipelines*

Streszczenie: W artykule przedstawiono badania wpływu wybranych czynników, w tym wstrząsów pochodzenia górniczego, opisanych za pomocą parametru PGV_{Hmax} na intensywność uszkodzeń przewodów podsystemu dystrybucji wody. Jako zmienne niezależne w utworzonych modelach regresji wielorakiej przyjęto: średnicę i materiał z którego wykonane są przewody, czas pracy bezuszkodzeniowej, wiek przewodów, wysokość ciśnienia i PGV_{Hmax}. Wartości PGV_{Hmax} w miejscach występowania awarii przewodów wodociągowych, wyznaczone zostały na podstawie zbudowanych modeli propagacji fali drgań w ośrodku gruntowym. Analiza przeprowadzona została dla próby losowej obejmującej sumarycznie wszystkie przewody sieci wodociągowej magistralnej, rozdzielczej i przyłącza oraz dla wyodrębnionych z tej grupy prób losowych obejmujących: przewody zbudowane ze stali i z żeliwa szarego, przewody z uwzględnieniem ich średnicy, średnicy i materiału oraz z uwzględnieniem czasu ich budowy.

Abstract: The article presents the influence of the selected factors, including mining tremors, described by the parameter PGV_{Hmax} on the failure rate of the water distribution pipelines. In created multiple regression models, the following independent variables were used: diameter and material from which the pipes were made, working time without failure, the age of pipes, the value of pressure and PGV_{Hmax}. The values of PGV_{Hmax} in places with damaged water pipelines were determined by constructed the seismic wave propagation models. The analysis was carried out for a random sample of all water pipelines and extracted from this sample new groups: steel and gray cast iron pipes, their diameters, diameters and materials, and their construction time.

1. Wprowadzenie

Systemy zaopatrzenia w wodę są częścią kluczowej, dla funkcjonowania gospodarki i społeczeństwa, infrastruktury technicznej. Do głównych zadań systemów zaopatrzenia w wodę należą: dostawa wody w wymaganej ilości, jakości oraz pod odpowiednim ciśnieniem [51]. Utrzymanie funkcjonalności i ciągłości działania tych systemów, wymaga zapewnienia ich odpowiedniego poziomu bezpieczeństwa [30,53,34,46], do czego zobowiązuje m. in. Dyrektywa Rady 2008/114/WE [8] oraz Ustawa z dnia 26 kwietnia 2007 r. o zarządzaniu kryzysowym [50]. Procedury ochrony infrastruktury krytycznej, do której zakwalifikowane zostało zaopatrzenie w wodę, jak również procedury wyznaczania niezawodności i bezpieczeństwa dostaw wody do odbiorców, obejmują w pierwszej

kolejności, identyfikację wszystkich czynników stwarzających zagrożenie [7,9,34,35,43,44,45].

W literaturze [5,15,17,29,35,40] spotkać można wiele analiz, dotyczących wpływu różnych czynników na intensywność uszkodzeń przewodów podsystemu dystrybucji wody. Wśród analizowanych czynników należy wymienić: średnicę, materiał, temperaturę otoczenia i wody w sieci wodociągowej, długość przewodów, średni opad deszczu, typ przewodu, głębokość ułożenia, czas pracy bezuszkodzeniowej, wysokość ciśnienia, rodzaj gruntu, agresywność medium i gruntu jak również wpływ eksploatacji górniczej, definiowany wyłącznie jako zmienna jakościowa (zero – jedynkowa). Wpływ wstrząsów górniczych na intensywność uszkodzeń przewodów podsystemu dystrybucji wody, nie był do chwili obecnej analizowany.

Zgodnie z powyższym, w celu uzupełnienia przeprowadzanych analiz awaryjności i niezawodności systemów zlokalizowanych na terenach górniczych, konieczne staje się uwzględnienie w analizach odpowiednio opisanego czynnika, informującego o wpływie wstrząsów górniczych.

Eksploatacja górnicza bardzo często prowadzona jest w Polsce pod obszarami silnie zurbanizowanym, gęsto uzbrojonymi w podziemną, liniową infrastrukturę techniczną tj. sieci: wodociągowe, kanalizacyjne, gazowe, ciepłownicze, teletechniczne. Stwarza to liczne zagrożenia zarówno dla powierzchniowych obiektów kubaturowych jak i dla sieci uzbrojenia podziemnego.

Sieci wodociągowe charakteryzują się różnymi rozwiązaniami konstrukcyjnymi, materiałowymi, sposobami połączeń, różny jest ich wiek i stan techniczny. Z tych właśnie powodów, przewody charakteryzuje zróżnicowana reakcja dynamiczna na dodatkowe obciążenia, wynikające z górniczo deformowanego terenu oraz z drgań powierzchniowych indukowanych wstrząsami górniczymi. Dodatkowo, bardzo trudna pozostaje identyfikacja stanu technicznego poszczególnych przewodów, czyli ich realnej wytrzymałości na obciążenia, co znacznie utrudnia przeprowadzanie zaawansowanych analiz.

Dla terenów sejsmicznych, opracowane zostały zależności opisujące wpływ trzęsień ziemi o określonych parametrach na intensywność uszkodzeń przewodów sieci wodociągowej. Przeprowadzane badania [23,39] wykazały, że PGV (Peak Ground Velocity - szczytowe wartości prędkości drgań gruntu) lepiej koreluje z uszkodzeniami rurociągów spowodowanymi trzęsieniami ziemi niż PGA (Peak Ground Acceleration - szczytowe wartości przyśpieszenia drgań gruntu). Wykazano, że PGV jest kierunkowo skorelowany z maksymalnymi odkształceniami gruntu – główną przyczyną uszkodzenia rurociągów podczas propagacji fali sejsmicznej. PGA lepiej koreluje z siłami bezwładności, które nie wpływają na obiekty podziemne. Pineda - Poras i Ordaz [42] wykazali, że model predykcji intensywności uszkodzeń przewodów sieci wodociągowej, oparty o parametr PGV, powoduje przeszacowanie liczby awarii powodowanych przez trzęsienia ziemi o magnitudzie 8,0 – 8,1 i mniejszej oraz niedoszacowanie liczby awarii dla silniejszych trzęsień ziemi. Parametrem, który lepiej opisuje zależność pomiędzy trzęsieniami ziemi a intensywnością uszkodzeń przewodów jest PGV^2/PGA.

Poza budową modeli empirycznych, przeprowadzonych zostało wiele innych, różnorodnych badań, dotyczących wpływu parametrów trzęsień ziemi na rurociągi zagłębione w gruncie. Przykładowo Wang and Cheng [52] wykazali poprzez budowę statycznego modelu numerycznego, że zachowanie rurociągu obciążonego trzęsieniem ziemi zależy od czasu przejścia fali jak również od niejednorodności gruntu, w którym zagłębiona jest rura. Takada i Tanabe [48] opracowali trójwymiarowy statyczny model numeryczny dla rurociągów oraz ich połączeń obciążonych dużymi trzęsieniami ziemi. Wykazano, że propagacja fal drgań w poszczególnych miejscach charakteryzowana jest przez szczytowe wartości parametrów równie dobrze jak przez prędkość propagacji fali. O'Rourke and Liu

[37] analizowali odkształcenia i krzywiznę gruntu powstające podczas propagacji fali trzęsienia ziemi, w zależności od różnych warunków gruntowych, na tworzenie się odkształceń powierzchni. Przejściowe odkształcenia powierzchni gruntu uważane są za najważniejsze z punktu widzenia wpływu propagacji fal sejsmicznych na podziemne obiekty liniowe tj. rurociągi i tunele. Lee [33] analizowała historię trzęsień ziemi i na jej podstawie badała wpływ parametrów trzęsień ziemi na zagłębiony gazociąg w gruncie z uwzględnieniem: typu materiału rurociągu (plastyczny, kruchy), średnicy, własności gruntu, głębokości ułożenia z uwzględnieniem lokalizacji (obszar mieszkalno przemysłowy, droga publiczna, tereny portowe - rurociąg zanurzony, itp.). O'Rourke and Ayala, [36] wykazali, że rurociągi wykonane z elastycznych materiałów są bardziej odporne na trzęsienia ziemi niż rurociągi kruche. Podobnie Bouziou and O'Rourke [4] wskazali że najbardziej odpornym na trzęsienie ziemi, które wystąpiło w dniu 22 lutego 2011 r. w miejscowości Christchurch materiałem przewodów wodociągowych był polichlorek winylu (PVC). O'Rourke i inni [38] analizowali zachowanie gruntu na płaszczyźnie styku z rurociągiem pod wpływem oddziaływania deformacji gruntu powodowanych trzęsieniami ziemi. Badania wykazały że wymagana jest zmiana założeń projektowych (obliczenia statyczno - wytrzymałościowe) pod kątem występujących sił od spodu rurociągu.

Wpływ trzęsień ziemi na podziemne, liniowe obiekty infrastruktury technicznej jest stosunkowo dobrze opisany w wielu pozycjach literaturowych na całym świecie. Problemem budowanych modeli propagacji fal sejsmicznych jest w dalszym ciągu dokładność modeli, na które wpływa duża różnorodność czynników, przyczyniających się do powstania określonego wpływu trzęsień ziemi na powierzchnię i obiekty.

Biorąc pod uwagę istotne różnice pomiędzy trzęsieniami ziemi a wstrząsami górniczymi, implementacja opracowanych dla trzęsień ziemi zależności, może być utrudniona i wymaga wykonania odpowiednich analiz. Nie zmienia to jednak faktu, że bogate opracowania dotyczące fal sejsmicznych stanowić mogą znakomite źródło wiedzy dla analiz wpływu wstrząsów górniczych na podziemne, liniowe obiekty techniczne.

Różnice pomiędzy trzęsieniami ziemi a wstrząsami górniczymi wynikają m in. z [55]:

- wielkości i intensywności – silne wstrząsy górnicze można porównać wyłącznie z małymi trzęsieniami ziemi,
- czasu trwania wymuszenia kinematycznego – czas trwania wstrząsów górniczych wynosi kilka sekund, podczas gdy trzęsienie ziemi może trwać 10-30 s,
- właściwości zapisów widmowych – co jest powodem znaczących różnic w oddziaływaniu zjawisk na obiekty budowlane,
- wartości szczytowych ruchu podłoża (PGA, PGV, PGD – trwała deformacja podłoża),
- różnic pomiędzy składowymi pionowymi i poziomymi drgań gruntu.

Stosunkowo prosto można zidentyfikować uszkodzenia rurociągów spowodowane trzęsieniami ziemi. Jest to możliwe dzięki skali oddziaływania zjawiska, jak również możliwości skorelowania jego wystąpienia ze wzrostem przepływu (na podstawie danych z monitoringu), informującym o wzroście ilości awarii systemu dystrybucji wody. Dodatkowo sama lokalizacja awarii może być prostsza, ze względu na wielkość i intensywność wypływu wody z uszkodzonego przewodu.

W przypadku awarii rurociągów, które mogą być spowodowane oddziaływaniem wstrząsów górniczych, sytuacja jest bardziej skomplikowana ponieważ może dotyczyć wyłącznie jednostkowych uszkodzeń, które z powodu niewystarczających badań, trudno wyodrębnić z grupy innych przyczyn powstania awarii.

W literaturze niejednokrotnie zwracano uwagę, że wpływ wstrząsów górniczych na obiekty liniowe zagłębione w gruncie nie jest wystarczająco poznany [20,21,22] lub

traktowany jest jako drugorzędny [28]. W pracy [22] wskazano że jednoznaczne określenie wpływu wstrząsów górniczych na uszkadzalność rurociągów było trudne, ze względu na możliwe przesunięcie czasowe przyczyny i skutków uszkodzeń. Pomimo tego, spotyka się publikacje przedstawiające wyniki badań (przede wszystkim analizy numeryczne lub statyczno - wytrzymałościowe), dotyczących wpływu wstrząsów górniczych na obiekty liniowe.

Dulińska [11] badała oddziaływanie drgań powierzchniowych na konstrukcję gazociągu, poprzez obliczenia odpowiedzi dynamicznej obiektu na silne wstrząsy. W analizach rozpatrywano 100 m odcinek gazociągu, uwzględniono 3 warianty gruntu: piaski suche, gliny ilaste i żwiry piaszczyste, założono sprężystą współpracę rury z gruntem w kierunku pionowym i poziomym. Badano model nierównomiernego wymuszenia kinematycznego (w każdym punkcie rurociągu w danej chwili działa inne wymuszenie kinematyczne). Analiza wykazała że większe odpowiedzi dynamiczne na wymuszenia kinematyczne uzyskiwane są dla mniejszych prędkości propagacji fali. Struktura częstotliwościowa drgań gruntu wpływa na wielkość odpowiedzi gazociągu, nawet w przypadku zbliżonych max amplitud drgań przy różnych strukturach częstotliwościowych. Max naprężenia powstałe w gazociągu pod wpływem działającego wstrząsu wyniosły 18 MPa czyli 10% wytrzymałości stali. W pracach [24] i [25] zwrócono uwagę na konieczność każdorazowego analizowania oddziaływania drgań podłoża na gazociągi oraz badanie ich odpowiedzi dynamicznej na zaistniałe drgania. Takie postępowanie, ma na celu zapobieganie możliwym konsekwencjom uszkodzeń starych stalowych gazociągów. Dulińska i Jasińska [12] badały za pomocą metod numerycznych zachowanie 100 m rurociągu stalowego ułożonego na betonowych podporach podczas wstrząsu sejsmicznego. Analizowane było rozchodzenie się fali drgań wzdłuż rurociągu. W rurociągu, w zależności od założonego modelu, obserwowane były, bądź nie, pęknięcia w miejscu lokalizacji podpór betonowych. Boron P., Dulińska J., [3] analizowali numeryczną odpowiedź połączeń (z uwzględnieniem śrub, kołnierzy i bloku oporowego) stalowych rurociągów na oddziaływanie wstrząsu sejsmicznego. Uzyskane odkształcenia śrub wskazały występowanie dużych wartości sił zginających, w miejscach połączeń rurociągu. W pracy [25] przedstawiona została analiza oddziaływania wstrząsu o energii $8 \cdot 10^7$ J na gazociąg stalowy DN200, niskiego ciśnienia, wyposażony w kompensatory dławikowe i charakteryzujący się dobrym stanem technicznym. W ramach pracy przeprowadzona została analiza statyczno – wytrzymałościowa i kinematyczna dla odcinka prostego gazociągu o długości 50 m oraz łuku 90^o. Przeprowadzona analiza wykazała że siły (przemieszczenia, naprężenia i odkształcenia) działające na gazociąg przyjmują wartości niższe od sił wywołanych przez ciągłe deformacje podłoża. Kurzeja J. [26] badała wpływ wstrząsów górniczych na węzeł autostrady A1 Piekary Śląskie. Analiza wskazała, że nie ma negatywnego wpływu wstrząsów górniczych na autostradę. Wnioskowanie przeprowadzone zostało na podstawie systematycznych obserwacji stanu autostrady.

Przedstawione w niniejszym artykule badania, maja na celu opracowanie zależności pomiędzy wstrząsami górniczymi, opisanymi za pomocą parametru PGV_{Hmax} a intensywnością uszkodzeń przewodów podsystemu dystrybucji wody z uwzględnieniem ich średnicy, materiału z którego zostały wykonane, czasu budowy, czasu pracy bezuszkodzeniowej oraz wysokości ciśnienia wody w przewodach. Są to pierwsze wyniki tego typu analiz publikowane w literaturze.

Opracowane zależności mogą posłużyć do prognozowania uszkodzeń przewodów sieci wodociągowych w analizowanym rejonie, znajdujących się pod wpływem oddziaływania obciążeń dynamicznych eksploatacji górniczej.

2. Metodyka badań

Przedmiotem badań był rzeczywisty podsystem dystrybucji wody, zlokalizowany w zasięgu wpływów naprężeń dynamicznych, spowodowanych wstrząsami pochodzenia górniczego. Przeprowadzone badania obejmowały analizę wszystkich awarii podsystemu dystrybucji wody oraz rejestracji wstrząsów górniczych o energii $E \geq 10^6$ J, występujących latach 2011-2014.

Zakres badań obejmował:

1. wyznaczenie lokalnej relacji tłumienia drgań gruntu dla wstrząsów górniczych opisanych za pomocą parametru PGV_{Hmax},
2. wyznaczenie wartości PGV_{Hmax} w miejscach występowania awarii przewodów wodociągowych,
3. budowę modeli regresji wielorakiej w celu identyfikacji czynników wpływających na intensywność uszkodzeń przewodów podsystemu dystrybucji wody.

2.1. Lokalna relacja tłumienia drgań gruntu dla wstrząsów górniczych opisanych za pomocą parametru PGV_{Hmax}

W literaturze [1,6,14,19,26], przedstawionych zostało wiele prób wyznaczania empirycznych zależności pomiędzy danym parametrem sejsmicznym a czynnikami wpływającymi na jego wielkość. Zależności te wyznaczane były zarówno dla wstrząsów górniczych jak i dla trzęsień ziemi.

W 1999 roku Si H. i Midorikawa S. [48] opracowały model regresji w którym jako czynniki wpływające na wielkość efektu sejsmicznego przyjęte zostały: energia wstrząsu, mechanizm ogniska, odległość od źródła wstrząsu oraz po raz pierwszy: budowa geologiczna ośrodka gruntowego.

Model posiada następującą postać:

$$logA = b - logR - k \cdot R \tag{1}$$

gdzie:

A – parametr sejsmiczny np.: PGV (Peak Ground Velocity - szczytowe wartości prędkości drgań gruntu), PGA (Peak Ground Acceleration - szczytowe wartości przyśpieszenia drgań gruntu) lub PGD (Peak Ground Displacement - szczytowe wartości przemieszczeń gruntu),

R – odległość od źródła wstrząsu, [m],

$$R = \sqrt{d^2 + h^2} \tag{2}$$

d – odległość epicentralna, [m],

h – pseudogłębokość (parametr zapewniający nieliniowość relacji dla małych odległości od źródła wstrząsu, dobrany tak aby błąd standardowy estymacji był jak najmniejszy), [m],

k – współczynnik przyjmujący następujące wartości: dla PGV – 0,002 dla PGA – 0,003, [-],

b – wskaźnik przesunięcia wyznaczany dla każdego zjawiska sejsmicznego z zależności, [-],

$$b = a \cdot \log E + c \cdot R + \sum d_i \cdot S_i + e + \varepsilon \tag{3}$$

gdzie:

a, c, d, e – współczynniki regresji estymowane z pomiarów sejsmometrycznych, [-],

E – energia wstrząsu, [J],

S_i – klasyfikator gruntu, parametr jakościowy, odległość d_i zależy od ilości przyjętych klas gruntu i, przyjęty w niniejszych badaniach zgodnie z Eurokod 8 (Tab. 1) [13], [-]

ε – parametr reprezentujący czynnik losowy, [-].

R – jak we wzorze (1)

Tab. 1. Klasyfikacja typu gruntu wg. Europejskiej Normy Eurokod 8 [13]. $V_{S,30}$ - średnia prędkość propagacji fal S w warstwach gruntu do głębokości 30 m.

Typ gruntu	**Opis stratygraficzny**	**$V_{S,30}$, m/s**
A	Skały lub inne skalne formacje geologiczne z co najwyżej pięciometrową słabą warstwą przypowierzchniową	>800
B	Piaski, żwiry i twarde gliny, co najmniej kilkunastometrowej miąższości, charakteryzujące się stopniowym wzrostem własności mechanicznych z głębokością	360-800
C	Gęste lub średnio gęste piaski, żwiry lub twarde gliny o miąższości od kilkunastu do kilkuset metrów	180-360
D	Utwory słabo lub średnio zwięzłe	<180
E	Profil gleby składający się z warstwy aluwium z wartością Vs odpowiadającą typom gruntu C i D oraz grubością 5-20m, leżącej na warstwie o Vs > 800 m/s	-

Parametrem przyjętym w niniejszej analizie do opisu drgań gruntu (parametr sejsmiczny) była maksymalna amplituda prędkości drgań poziomych, wyznaczona jako poziome maksimum długości wektora drgań gruntu (PGV_{Hmax}).

Model (1) stał się podstawą do opracowania relacji tłumienia drgań gruntu dla rejonu, w którym usytuowany jest badany podsystem dystrybucji wody oraz do wyznaczenia wartości parametru sejsmicznego w miejscach występowania awarii wodociągowych.

2.2. Wyznaczenie wartości PGV_{Hmax} w miejscach występowania awarii przewodów wodociągowych

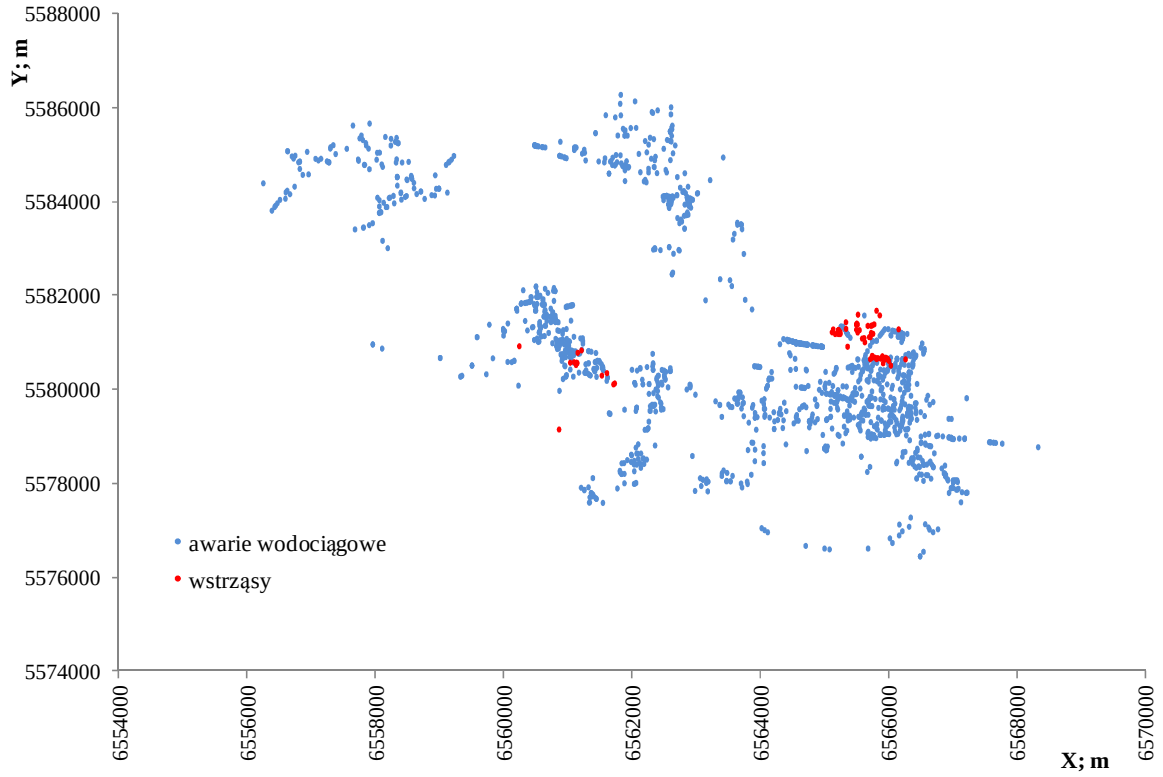
Procedura wymagała przyporządkowania poszczególnym wstrząsom górniczym powodowanych przez nie awarii przewodów wodociągowych. W tym celu konieczne stało się przyjęcie następującego założenia: wstrząs górniczy może być przyczyną każdej awarii wodociągowej występującej do 30 dni po wstrząsie.

Konieczność przyjęcia tego założenia, spowodowana była rodzajem posiadanych przez przedsiębiorstwo wodociągowe danych o awariach przewodów wodociągowych. Analizowana baza danych posiadała wyłącznie zestawienia czasu rozpoczęcia prac budowlanych związanych z usuwaniem poszczególnych awarii. Dla celów badań ważne było uzyskanie informacji o czasie zaistnienia awarii. Moment wystąpienia awarii ze względu na dużą liczbę niewiadomych oraz małą dokładność urządzeń pomiarowych nie jest możliwy w chwili obecnej do dokładnego ustalenia. Zgodnie z powyższym założone 30 dni stanowi pewnego rodzaju bufor, w którym dana awaria mogła się rozpocząć.

W badaniach nie zostało przyjęte żadne przestrzenne ograniczenie dotyczące maksymalnej odległości, pomiędzy danym wstrząsem, a awarią przewodu wodociągowego, w związku z nieznanym efektywnym zasięgiem oddziaływania wstrząsów górniczych na podziemne obiekty liniowe.

Wyznaczenie wartości PGV_{Hmax} w miejscu występowania awarii sieci wodociągowej, z wykorzystaniem modelu powstałego na podstawie równania (1), wymagało: przeprowadzenia klasyfikacji gruntu w miejscu występowania poszczególnych awarii zgodnie z Eurokod 8 [13] oraz wyznaczenia odległości epicentralnej (pomiędzy wstrząsem górniczym a awarią wodociągową) w oparciu o posiadane współrzędne geograficzne.

W przypadku gdy ta sama awaria wodociągowa, przypisana została do różnych wstrząsów górniczych, w analizie pozostawiano tą parę „wstrząs – awaria”, w której parametr PGV_{Hmax} przyjmował wyższą wartość. Ostatecznie otrzymano 993 pary „wstrząs – awaria”. Na Rys. 1 przedstawiono przestrzenne rozmieszczenie wszystkich analizowanych wstrząsów górniczych i awarii podsystemu dystrybucji wody.



Rys. 1. Wstrząsy górnicze o $E \geq 10^6$J i awarie sieci wodociągowej występujące w latach 2011 – 2014.

2.3. Wpływ wybranych czynników na intensywność uszkodzeń przewodów podsystemu dystrybucji wody

W kolejnym etapie badań dla poszczególnych przewodów wodociągowych, scharakteryzowanych za pomocą parametrów: liczba awarii, średnica, materiał, wiek itp. przypisano średnią wartość PGV_{Hmax} ze wszystkich miejsc występowania awarii wodociągowych na danym przewodzie. W badaniach, poprzez przewód wodociągowy, rozumiano odcinki sieci wodociągowej magistralnej i rozdzielczej oraz przyłącza, posiadające jednorodne parametry w zakresie: średnicy, materiału oraz czasu budowy. Analizę wpływu wybranych czynników na intensywność uszkodzeń przewodów podsystemu dystrybucji wody przeprowadzono dla grupy 792 przewodów.

W modelu regresji wielorakiej, wpływu wybranych czynników na jednostkową intensywność uszkodzeń przewodów podsystemu dystrybucji wody, przyjęto następujące zmienne:

- jednostkowa intensywność uszkodzeń przewodów (λ) – zmienna zależna budowanych modeli regresji,
  oraz zmienne niezależne:
- materiał (M) i średnica poszczególnych przewodów (DN),
- średnia wysokość ciśnienia w miejscu występowania awarii na danym przewodzie (P),
- średni wiek przewodu w momencie uszkodzenia (W),

- czas pracy bezuszkodzeniowej przewodu w analizowanym czasie (Tp),
- PGV_{Hmax},

Najczęściej stosowanymi parametrami do oceny awaryjności przewodów wodociągowych są: strumień uszkodzeń ω(t) lub intensywność uszkodzeń λ(t). Przyjmując założenie, że strumień uszkodzeń jest strumieniem: bez następstw (pojawiające się w poszczególnych czasach Δt uszkodzenia są niezależne), pojedynczym (prawdopodobieństwo wystąpienia więcej niż jednego uszkodzenia na dowolnie małym odcinku czasu dąży do zera) i stacjonarnym (prawdopodobieństwo zaistnienia k uszkodzeń w przedziale czasu Δt zależy tylko od tego przedziału, a nie zależy od położenia na osi czasu), otrzymuje się równość tego parametru z intensywnością uszkodzeń λ(t) = ω(t) = const [50,52]. W oparciu o przedstawione założenie, w przeprowadzonych badaniach awaryjność przewodów wodociągowych opisana została za pomocą wskaźnika intensywności uszkodzeń.
Jednostkowa intensywność uszkodzeń λ(Δt) dla elementów liniowych (rurociągi) wyraża równanie [31,53]:

$$\lambda(\Delta t) = \frac{n(\Delta t)}{L\,\Delta t}\ ;\ [\text{uszk./(km}\cdot\text{a)}] \qquad (4)$$

gdzie:
λ(Δt) – jednostkowa intensywność uszkodzeń obiektów liniowych, [uszk./(km· a)],
n(Δt) – liczba uszkodzeń w przedziale czasu Δt,
L – średnia długość badanych rurociągów w przedziale czasu Δt, [km],
Δt – przedział czasu na jaki podzielono okres obserwacji, [a].

Intensywność uszkodzeń analizowanych przewodów sieci wodociągowej wynosiła λmaxs = 200 uszk./(km·4lata), mediana = 8,8 uszk./(km·4lata), a intensywność uszkodzeń przyłączy λmaxp = 2000 uszk./(km·4lata), mediana = 93,9 uszk./(km·4lata).

System zaopatrzenia w wodę można traktować jako odnawialny (występujące cyklicznie stany pracy i niesprawności), z punktu widzenia teorii niezawodności. Nazywa się on „modelem dwustanowym z odnową niezerową” [31,53]. W przypadku gdy czas pracy bezuszkodzeniowej ma rozkład wykładniczy:

$$T_p = \frac{1}{\lambda} \qquad (5)$$

gdzie:
Tp – średni czas pracy bezuszkodzeniowej elementu, [lata]
λ – intensywność uszkodzeń obiektów liniowych, [uszk./(km rok)],

W przeprowadzonej analizie czas pracy bezuszkodzeniowej zawierał się w przedziale (26 – 730) dni. Należy zaznaczyć, że większość przewodów (w szczególności przyłącza) ulegała uszkodzeniu jeden raz w analizowanym czasie, czas pracy bezuszkodzeniowej przyjmował wtedy tą samą wartość równą 730 dni (mediana).

Wysokość ciśnienia w miejscu występowania awarii wodociągowych, obliczona została na podstawie znajomości (odczyt z wykresów monitoringu) wysokości ciśnienia w punkcie zasilającym dany przewód lub rejon sieci wodociągowej, w której usytuowany był dany przewód oraz różnicy wysokości pomiędzy punktem zasilającym a miejscem wystąpienia awarii, odczytanej z numerycznego modelu terenu. Dla poszczególnych przewodów, wysokość ciśnienia wyznaczona została jako średnia arytmetyczna wysokości ciśnienia ze wszystkich miejsc występowania awarii na danym przewodzie.

Wysokość ciśnienia dla wszystkich analizowanych przewodów zawierała się w przedziale (13,39 - 86,87) m. sł. w., podczas gdy mediana wynosiła: 44,48 m. sł. w.

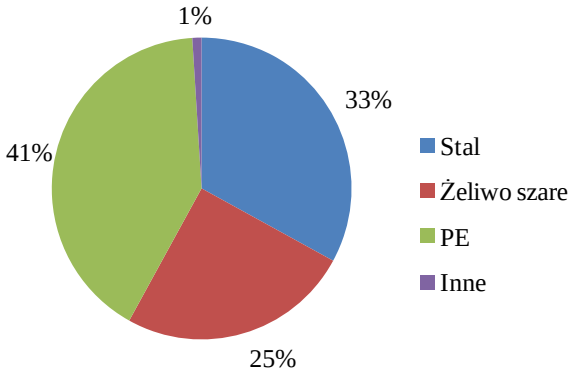
Modele regresji, opracowane zostały dla próby losowej składającej się z: przewodów sieci wodociągowej magistralnej, rozdzielczej i przyłączy oraz dla wyodrębnionych z tej próby losowej grup przewodów: przewody zbudowane ze stali, przewody zbudowane z żeliwa szarego, przewody o średnicach: DN80, DN100, DN150, DN100 zbudowane ze stali, DN100 zbudowane z żeliwa szarego, DN150 zbudowane ze stali, przewody wybudowane w następujących przedziałach czasu: 1885-1920, 1922-1938, 1939-1961, 1962-1978, 1979-1994, 1995-2012.

Analiza wykonana została w pakiecie oprogramowania Statistica (Statsoft) z wykorzystaniem regresji wielorakiej.

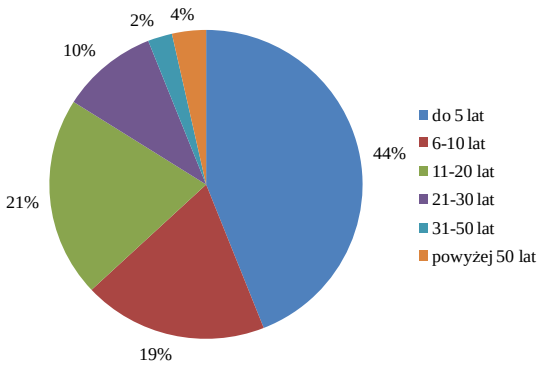
2.4. Charakterystyka badanego podsystemu dystrybucji wody

Badany podsystem dystrybucji wody zlokalizowany jest na Górnym Śląsku i odpowiada za dostawę wody do ok. 160 tyś. mieszkańców gminy (163 tyś. mieszkańców w 2011 r. i 157 tyś. w 2014 r.). Źródłem zaopatrzenia w wodę podsystemu dystrybucji są ujęcia wód podziemnych i powierzchniowych, których właścicielem jest Górnośląskie Przedsiębiorstwo Wodociągów S. A. (GPW). Zakup wody od GPW odbywa się poprzez 56 studni zakupowych i średnio w analizowanych latach 2011 – 2014 wynosił 7,3 mln m^3/ rok, podczas gdy sprzedaż wody wyniosła średnio ok. 6,2 mln. m^3/rok.

Sieć wodociągowa w zakresie średnic DN20 – DN600 ma ok. 450 km długości w tym ok. 200 km położonych jest w zasięgu oddziaływania eksploatacji górniczej. Przewody sieci wodociągowej wykonane są przede wszystkim z: polietylenu, stali i żeliwa (Rys. 2), i eksploatowane są w większości (60%) przez czas krótszy niż 10 lat (Rys. 3).



Rys. 2. Struktura materiałowa sieci wodociągowej w roku 2014. Na wykresie „Inne” – azbestocement i PVC.



Rys. 3. Struktura wiekowa sieci wodociągowej w roku 2014.

Wartość wskaźnika intensywności uszkodzeń dla badanego podsystemu dystrybucji wody w poszczególnych latach (Tab. 2) świadczy o dużej awaryjności przewodów wodociągowych $\lambda > 0,5$ uszk./(km·rok).

Tab. 2. Zestawienie wartości jednostkowej intensywności uszkodzeń w latach 2011 – 2014 dla badanego podsystemu dystrybucji wody.

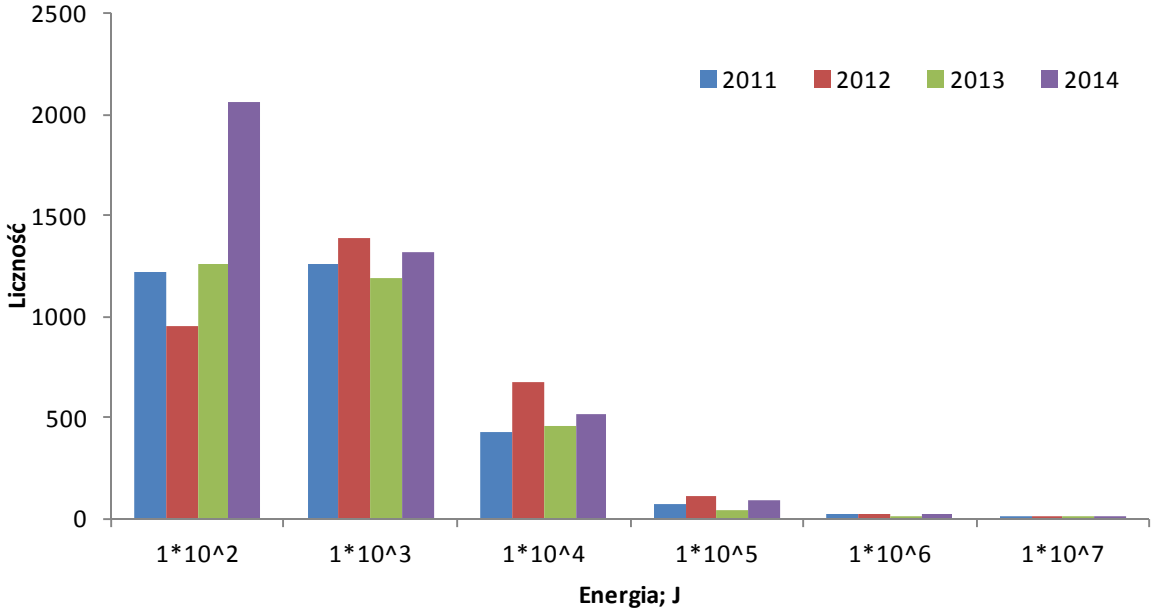
Rok	**Liczba awarii**	**Długość sieci; km**	**Jednostkowa intensywność uszkodzeń λ; uszk./(km·rok)**
2011	445	443,9	1,00
2012	532	444,9	1,20

2013	299	456,4	0,66
2014	285	454,6	0,63

2.5. Charakterystyka wstrząsów górniczych występujących w analizowanym rejonie

Badany podsystem dystrybucji wody zlokalizowany jest na terenie gminy, w której od XIX wieku prowadzona jest aktywna eksploatacja górnicza.

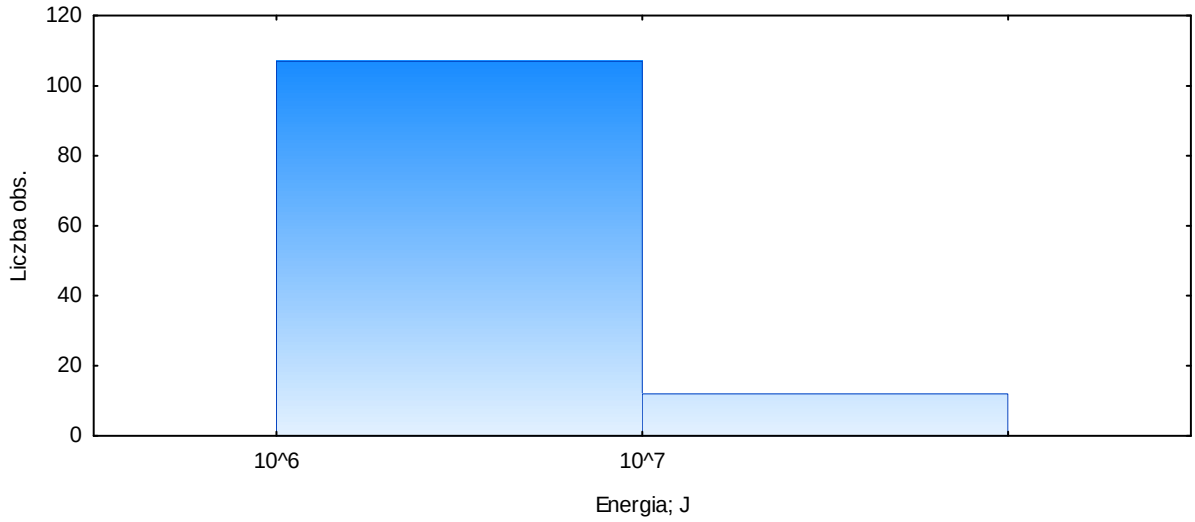
W chwili obecnej wstrząsy górnicze rejestrowane są na 7 stanowiskach pomiarowych wyposażonych w aparaturę Amax-GSI, produkowaną w Głównym Instytucie Górnictwa. Aparatura Amax-GSI [41] jest wielokanałową aparaturą specjalistyczną, służącą do bezpośredniego pomiaru drgań na powierzchni. Rejestracja przyśpieszeń drgań gruntu odbywa się za pomocą rejestratora i trójskładowego zespołu akcelerometrów piezoelektrycznych. Dzięki możliwości scałkowania sygnału, uzyskuje się zapis prędkości drgań w trzech prostopadłych płaszczyznach. Zakres energetyczny rejestrowanych na analizowanym rejonie wstrząsów wynosi od $1 \cdot 10^2$ do $1 \cdot 10^7$ J (Rys. 4).



Rys. 4. Zestawienie liczby wstrząsów rejestrowanych na obszarze górniczym badanej kopalni w zakresie energetycznym od $1 \cdot 10^2$ do $1 \cdot 10^7$ J w latach 2011-2014.

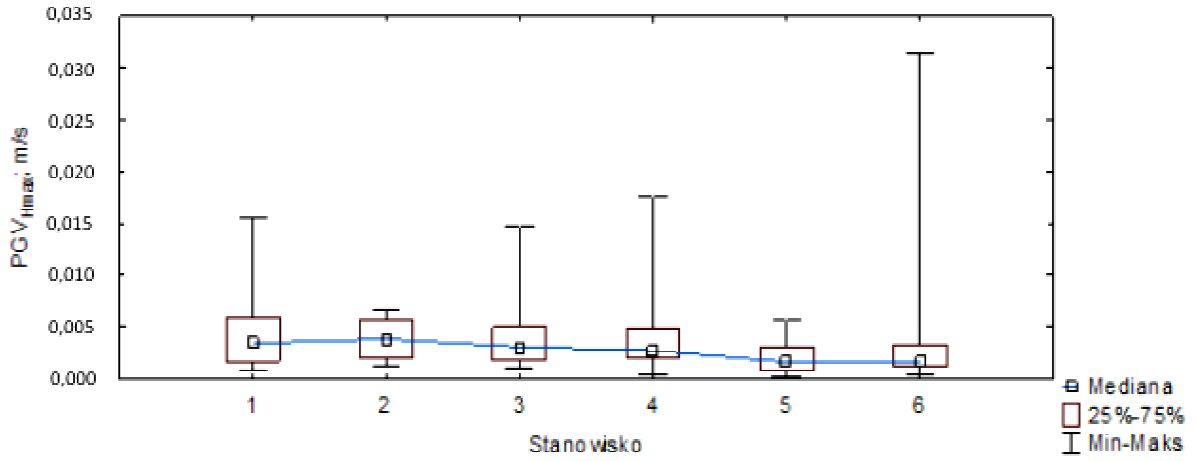
W analizowanym rejonie, najwięcej występuje wstrząsów słabych, nieodczuwalnych przez człowieka ($E < 10^4$ J) i rejestrowanych wyłącznie przez specjalistyczną aparaturę.

W ramach badań analizowano wyłącznie wstrząsy o $E \geq 10^6$ J rejestrowane na 6 stanowiskach pomiarowych ze względu na dużą liczbę niekompletnych zapisów na jednym ze stanowisk. Ostatecznie do analizy przyjęto 113 wstrząsów górniczych (252 rejestracji), których zestawienie energetyczne przedstawiono na Rys. 5.



Rys. 5. Histogram energii wstrząsów występujących w latach 2010 – 2014.

Na Rys. 6 przedstawiono wartości składowej poziomej prędkości drgań, rejestrowane na poszczególnych stanowiskach pomiarowych. Zakres rejestrowanych wartości wynosi od 0,0002 do 0,3150 m/s, podczas gdy mediana dla wszystkich stanowisk jest podobna i wynosi ok. 0,0030 m/s. Największa wartość prędkości drgań gruntu zarejestrowana została 16.02.2013 r. na 6 stanowisku pomiarowym natomiast najmniejsze wartości na stanowiskach 5 i 2.



Rys. 6. Rozkład median wartości składowej poziomej szczytowych prędkości drgań gruntu dla poszczególnych stanowisk pomiarowych dla wstrząsów o energii $E \geq 10^6$ J występujących w latach 2010 – 2014.

3. Lokalna relacja tłumienia drgań gruntu

Estymację poszczególnych parametrów równania regresji przeprowadzono z wykorzystaniem metody najmniejszych kwadratów. Z modelu wyeliminowane zostały zmienne niezależne, które charakteryzowały się najmniejszą wartością stosunku wariancji wyjaśnionej do niewyjaśnionej. Ostatecznie lokalna relacja tłumienia drgań gruntu dla $\log PGV_{Hmax}$, dla poziomu istotności $\lambda = 0{,}05$ przyjęła następującą postać:

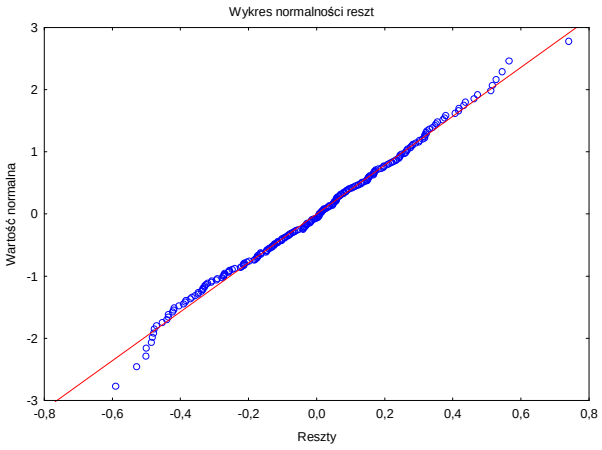
$$\log PGV_{Hmax} = 0{,}44937 \cdot \log E - 1{,}49304 \cdot \log R + 0{,}07623 \cdot S_i - 0{,}88661 \quad (8)$$

gdzie:
PGV_{Hmax} – składowa pozioma szczytowej wartości prędkości drgań gruntu; [m/s],
E – energia wstrząsu; [J],
R – odległość od źródła wstrząsu, dla pseudogłębokości h = 1022m dla której wartość błędu standardowego estymacji zgodnie z równaniem (3) przyjmuje najmniejszą wartość, [m],
S_i – zmienna jakościowa przyjmująca wartość $S_i = 0$ dla gruntu typu B i $S_i = 1$ dla gruntu typu C; [-].

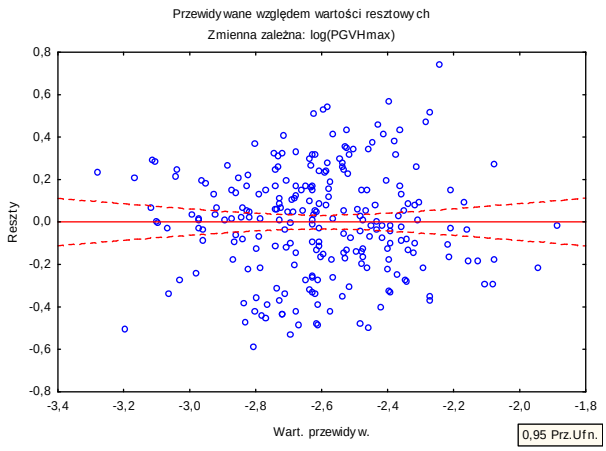
Współczynnik korelacji dla opracowanego modelu wynosi 0,68840. Wartość współczynnika determinacji R^2 = 0,47390, oznacza umiarkowane dopasowanie modelu do analizowanych danych. Na wartość otrzymanego współczynnika determinacji złożyły się m. in.: niedokładność wyznaczania energii i ognisk wstrząsów [2], nieuwzględnienie w modelach zróżnicowania radiacji energii z ogniska [2], nie rozpoznana w pełni budowa geologiczna ośrodka gruntowego [2], błędy pomiaru akcelometrów, błędy modelu regresji. W większości spotykanych w literaturze modeli propagacji fal sejsmicznych [2,6,10,26,27] współczynnik determinacji przyjmuje podobne wartości. Zgodnie z powyższym, postanowiono kontynuować analizę w oparciu o otrzymany model regresji.

Standardowy błąd estymacji modelu wynosi Se = 0,25344 i oznacza, że szacowane wartości zmiennej $\log PGV_{Hmax}$ różnią się od wartości empirycznych średnio o 0,25344 m/s.

Budowę modelu regresji wielorakiej zakończono analizą reszt, która potwierdziła spełnienie przez model założeń klasycznej metody najmniejszych kwadratów.



Rys. 7. Wykres normalności reszt modelu predykcji zmiennej $\log PGV_{Hmax}$.



Rys. 8. Wykres rozrzutu reszt modelu predykcji zmiennej $\log PGV_{Hmax}$.

4. Wpływ wybranych czynników na intensywność uszkodzeń podsystemu dystrybucji wody

Wyniki analizy intensywności uszkodzeń wszystkich przewodów sieci wodociągowej oraz grup: przewodów stalowych i wykonanych z żeliwa szarego przedstawiono w Tab. 3.

Tab. 3. Zestawienie modeli regresji dla intensywności uszkodzeń przewodów wodociągowych.

Rodzaj przewodów	**Model regresji**	**R, R^2, Se**

Wszystkie przewody	$\log\lambda spv = -0{,}03547 \cdot DN + 0{,}68996 \cdot M + 73{,}53762 \cdot PGV_{Hmax} + 0{,}00085 \cdot Tp + 1{,}62527$	R= 0,84113 R^2= 0,70750 Se = 1,11430
Przewody stalowe	$\log\lambda spvs = -0{,}04067 \cdot DN + 72{,}90571 \cdot PGV_{Hmax} + 3{,}17737$	R= 0,85774 R^2= 0,73572 Se = 1,03300
Przewody z żeliwa szarego	$\log\lambda spvz = -0{,}00297 \cdot DN - 0{,}00068 \cdot Tp - 1{,}42276$	R= 0,69170 R^2= 0,47845 Se = 0,26376

W przypadku grupy wszystkich przewodów sieci wodociągowej oraz przewodów stalowych, otrzymane wartości współczynników determinacji, świadczą o dobrym dopasowaniu modeli do danych. Parametr PGV_{Hmax} informujący o oddziaływaniu wstrząsów górniczych, okazał się istotny w przypadku tych modeli i posiada dodatni wpływ na zmienną zależną. Największy wpływ na zmienną zależną w modelu dla wszystkich przewodów wywierają kolejno: DN i M, Tp i PGV_{Hmax}, a w przypadku przewodów stalowych: DN i PGV_{Hmax}. Przykładowo dla modelu zbudowanego dla przewodów stalowych w przypadku wzrostu PGV_{Hmax} o wartość 0,01 m/s, zmienna zależna $\log\lambda spvs$ wzrośnie o 0,72906 (intensywność uszkodzeń wzrośnie o 5,35797 uszk./(km·4lata)).

Dla przewodów wykonanych z żeliwa szarego, otrzymano słabszą zależność (R^2= 0,47845) a zmienna informująca o oddziaływaniu wstrząsów górniczych na przewody wodociągowe okazała się nieistotna.

W całym zakresie analizowanych średnic przewodów (DN20 – DN600), tylko dla DN80, DN100 i DN150 udało się otrzymać zależności funkcyjne, pomiędzy wybranymi czynnikami a intensywnością uszkodzeń przewodów. Dla pozostałych zmiennych, albo brakowało wystarczającej liczby danych albo zmienne niezależne okazywały się nieistotne albo współczynniki korelacji i determinacji wynosiły zero. Zestawienie otrzymanych modeli regresji wielorakiej zawiera Tab. 3.

Tab. 4. Zestawienie modeli regresji dla intensywności uszkodzeń przewodów wodociągowych o średnicach DN80, DN100 i DN150.

Średnica	**Model regresji**	**R, R^2, Se**
DN80	$\log\lambda sv80 = -0{,}01010 \cdot W - 1{,}20001$	R= 0,53986 R^2= 0,29144 Se = 0,39376
DN100	$\log\lambda sv100 = -0{,}03320 \cdot W + 67{,}93168 \cdot PGV_{Hmax} - 0{,}00070 \cdot Tp - 1{,}60510$	R= 0,64731 R^2= 0,41902 Se = 0,26561
DN150	$\log\lambda sv150 = -0{,}00113 \cdot Tp - 1{,}55594$	R= 0,66583 R^2= 0,44333 Se = 0,29573

Tylko w modelu zbudowanym dla przewodów o średnicy DN100, zmienna opisująca wpływ wstrząsów górniczych okazała się istotna. Model wyjaśnia ok. 40% zmienności zmiennej zależnej. Współczynnik korelacji jest bliski wartości 0,65. Największy wpływ na

zmienną zależną posiadają kolejno: Tp, PGV_{Hmax} i W. Modele zbudowane dla DN80 i DN150, posiadają niskie wartości współczynników determinacji.

W przypadku analizy intensywności uszkodzeń z uwzględnieniem średnic i materiałów, z których wykonane zostały przewody, udało się zbudować modele regresji wyłącznie dla przewodów: DN100 wykonanych ze stali, DN100 wykonanych z żeliwa szarego i DN150 wykonanych ze stali. Wyniki przeprowadzanych analiz zestawiono w Tab. 4.

Tab. 5. Zestawienie modeli regresji dla intensywności uszkodzeń przewodów wodociągowych poszczególnych średnic i materiałów.

Średnica i materiał	**Model regresji**	**R, R^2, Se**
DN100 stal	$\log\lambda sv100s = 64{,}44562 \cdot PGV_{Hmax} - 0{,}00083 \cdot Tp - 1{,}70361$	R= 0,68320 R^2= 0,46677 Se = 0,25054
DN100 żeliwo szare	$\log\lambda sv100z = - 0{,}00088 \cdot Tp - 1{,}63868$	R= 0,49176 R^2= 0,24183 Se = 0,27936
DN150 stal	$\log\lambda sv150s = - 0{,}00135 \cdot Tp - 1{,}45395$	R= 0,74349 R^2= 0,55278 Se = 0,28652

Tylko w modelu dla przewodów wykonanych ze stali o średnicy DN100, zmienna opisująca wpływ wstrząsów górniczych (PGV_{Hmax}) okazała się zmienną istotną. Zmienna posiada dodatni wpływ na intensywność uszkodzeń przedmiotowych przewodów. W modelu również Tp oraz wyraz wolny mają istotny wpływ na zmienną zależną. Model wyjaśnia ok. 47% zmienności zmiennej $\log\lambda sv100s$.

Dla przewodów wykonanych z żeliwa szarego o średnicy DN100 oraz dla przewodów wykonanych ze stali o średnicy DN150, tylko Tp oraz wyraz wolny okazały się zmiennymi istotnymi. Otrzymana zależność dla przewodów żeliwnych DN100 jest bardzo słaba, wartość współczynnika determinacji wynosi zaledwie 25% i świadczy o braku zależności. Lepszy model udało się uzyskać dla przewodów stalowych DN150, w którym współczynnik determinacji wynosi ok. 55%.

W związku z tym że przeprowadzane badania obejmowały przewody wodociągowe w wieku nawet 129 lat, uzasadnionym wydaje się przeprowadzenie analiz awaryjności przewodów wykonanych w poszczególnych przedziałach czasu (Tab. 6).

Tab. 6. Modele regresji dla intensywności uszkodzeń przewodów sieci wodociągowej i przyłączy wybudowanych w różnych okresach czasu.

Grupa	**Lata**	**Model regresji**	**R, R^2, Se**
1	1885-1920	$\log\lambda_{spv(1885\text{-}1920)} = - 0{,}03320 \cdot DN + 1{,}34550 \cdot M + 216{,}44430 \cdot PGV_{Hmax} + 0{,}03000 \cdot P$	R= 0,90274 R^2= 0,81494 Se = 1,06430
2	1922-1938	$\log\lambda_{spv(1922\text{-}1938)} = - 0{,}01687 \cdot DN + 1{,}61912$	R= 0,73161 R^2= 0,53525

			Se = 1,57350
3	1939-1961	$\log\lambda_{spv(1939\text{-}1961)} = -0{,}04136 \cdot DN + 1{,}08442 \cdot M + 2{,}33101$	R= 0,86648 R^2= 0,75078 Se = 1,09150
4	1962-1978	$\log\lambda_{spv(1962\text{-}1978)} = -0{,}03291 \cdot DN + 0{,}62525 \cdot M + 50{,}00144 \cdot Tp + 1{,}28402$	R= 0,86310 R^2= 0,74495 Se = 0,99725
5	1979-1994	$\log\lambda_{spv(1979\text{-}1994)} = -0{,}04550 \cdot DN + 3{,}65622$	R= 0,87887 R^2= 0,77242 Se = 0,96855
6	1995-2012	$\log\lambda_{spv(1995\text{-}2012)} = -0{,}050443 \cdot DN - 0{,}645707 \cdot M + 4{,}442355$	R= 0,90993 R^2= 0,82797 Se = 0,73568

Otrzymane modele regresji cechują się bardzo dobrym dopasowaniem do obserwacji, współczynnik determinacji zawiera się w większości przypadków w przedziale od 74 do 83%. Tylko model dla przewodów wybudowanych w latach 1922 – 1938 posiada współczynnik determinacji wynoszący ok. 53%. Zmienna charakteryzująca wstrząsy górnicze, okazała się istotna tylko dla modelu zbudowanego dla przewodów wybudowanych w latach 1885 – 1920. Zgodnie z powyższym, można wnioskować, że najstarsze analizowane przewody, mogą być najmniej odporne na występujące wstrząsy górnicze. W przedstawionym modelu, istotną zmienną okazała się również wysokość ciśnienia w miejscu występowania awarii, co może również świadczyć o mniejszej odporności wyłącznie najstarszych przewodów na wymieniony czynnik. Dla przewodów zbudowanych w latach 1922 – 1938 i 1979 – 1994, wyłącznie średnica z wyrazem wolnym, wywierały istotny wpływ na intensywność uszkodzeń przewodów, co jest najprawdopodobniej skutkiem znaczących różnic w sposobie budowy/posadowienia przewodów poszczególnych średnic.

Dla pozostałych grup przewodów, analizowanych ze względu na czas ich budowy, parametrami istotnymi w tworzonych modelach okazały się: średnica i materiał przewodów, czas pracy bezuszkodzeniowej oraz wyraz wolny.

Przeprowadzona analiza reszt, dla każdego omówionego powyżej modelu regresji, wykazała, że spełnione są założenia analizy regresji w zakresie normalności reszt, liniowości modelu względem parametrów oraz homoscedastyczności dla modeli zbudowanych dla następujących zmiennych: $\log\lambda spvz$, $\log\lambda sv80$, $\log\lambda sv100$, $\log\lambda sv150$, $\log\lambda sv100s$, $\log\lambda sv100z$, $\log\lambda sv150s$. W przypadku pozostałych modeli warunki o liniowości modelu względem parametrów oraz o homoscedastyczności nie zostały spełnione. Na otrzymanych wykresach rozrzutu widoczne są dwie grupy punktów, ze względu na dużą rozbieżność pomiędzy wartościami intensywności uszkodzeń sieci wodociągowej i przyłączy. Opisana sytuacja świadczy o niejednorodności próby losowej, przyjętej do analizy i może wpływać na jakość otrzymywanych modeli. Nie zmienia to jednak faktu, że wartości współczynników determinacji otrzymanych w przedmiotowych modelach, ze względu na większą liczbę danych przy uwzględnianiu całej grupy przewodów, przyjmują dobre wartości. Zgodnie z powyższym uzasadnione wydaje się przeprowadzenie podobnej analizy indywidualnie dla przewodów sieci wodociągowej i przyłączy.

5. Podsumowanie i wnioski

Przedstawione w artykule analizy, stanowią pierwszą tego rodzaju próbę, wyznaczenia wpływu wstrząsów górniczych na intensywność uszkodzeń przewodów podsystemu dystrybucji wody.

Przeprowadzone badania opierały się wyłącznie na danych historycznych otrzymanych z obiektu rzeczywistego, co uniemożliwiło powtórzenie i eliminację pewnych błędów pomiaru. Należy również podkreślić, że badania obejmowały teren całej gminy o powierzchni 69,44 km^2. Powoduje to możliwość występowania wielu dodatkowych, nie uwzględnionych/niezidentyfikowanych w przeprowadzonych badaniach czynników, które mogły wpływać na zmienne zależne a zatem także, na jakość otrzymywanych modeli.

Analizy przedstawione w artykule pozwoliły na sformułowanie następujących wniosków:

1. Zmienna niezależna (PGV_{Hmax}) informująca o wpływie wstrząsów górniczych na intensywność uszkodzeń przewodów podsystemu dystrybucji wody, okazała się istotna w zbudowanych modelach regresji dla wybranych grup przewodów podsystemu dystrybucji wody. Wskazuje to na istnienie nieuwzględnianego do tej pory czynnika wpływającego na awaryjność przewodów wodociągowych, zlokalizowanych na terenach górniczych podatnych na tąpania.
2. Zmienną niezależną, która w większości zbudowanych modeli regresji posiadała największy wpływ na zmienną zależną okazała się średnica przewodów (9 z 15 modeli) oraz kolejno czas pracy bezuszkodzeniowej (7 z 15 modeli) oraz PGV_{Hmax} i materiał z którego wykonane są przewody (po 5 z 15 modeli).
3. Opracowana w pracy metodyka analizy i oceny wpływu wstrząsów górniczych na intensywność uszkodzeń przewodów podsystemu dystrybucji wody, może zostać zastosowana dla każdego podsystemu dystrybucji wody oraz innych ciśnieniowych, liniowych obiektów podziemnych tj. sieci ciepłownicze, gazowe, kanalizacja ciśnieniowa, zlokalizowanych na terenach oddziaływania wstrząsów górniczych, pod warunkiem, że dostępne będą dane dotyczące: awaryjności obiektów liniowych, klasy gruntów, na których zlokalizowane są obiekty, wartości energii wstrząsów, współrzędnych epicentrum wstrząsów oraz postaci modeli rozprzestrzeniania się fali drgań gruntu w rozpatrywanym rejonie lub w przypadku braku takich modeli - wartości PGV_{Hmax}, jak również współrzędnych stanowisk pomiarowych i typu gruntu, na którym są one zlokalizowane.
4. Opracowane modele regresji mogą stanowić podstawę do opracowania programu komputerowego, który mógłby posłużyć do prognozowania awarii przewodów podsystemu dystrybucji wody po wstrząsie górniczym, o określonych parametrach np. bazując na prognozie wstrząsów górniczych, podczas rozpoczynania eksploatacji w nowym polu wydobywczym.

Literatura

1. Atkinson G. M., Boor D. M., Earthquake ground-motion prediction equation for Western North America, Bulletin of the Seismological Society of America 2006; 96: 2181-2205, https://doi.org/10.1785/0120050245.
2. Bańka P., Kołodziejczyk P., Lier E., Wykorzystanie wyników pomiarów parametrów drgań gruntu do wyznaczenia wartości współczynnika amplifikacji drgań, Przegląd Górniczy 2016, 4 (1121): 71-79.
3. Boron P., Dulińska J., The dynamic analysis of a steel pipeline under a seismic shock, Procedia Engineering 2017, 199: 104 – 109, https://doi.org/10.1016/j.proeng.2017.09.166.

4. Bouziou D., O'Rourke T. D., Response of the Christchurch water distribution system to the 22 February 2011 earthquake, Soil Dynamics and Earthquake Engineering 2017, 97: 14 – 24, https://doi.org/10.1016/j.soildyn.2017.01.035.
5. Bubtiena A. M., El Shafie A. H., Jaafar O., Performance improvement for pipe breakage prediction modeling using regression method, International Journal of the Physical Sciences 2011, 25 (6): 6025 – 6035, http://www.academicjournals.org/journal/IJPS/article-full-text-pdf/F7B3D0625658.
6. Chodacki J., New ground motion prediction equation for peak ground velocity and duration of ground motion for mining tremors in Upper Silesia, Acta Geophysica 2016, 64 (6): 2449 – 2470, http://agp.igf.edu.pl/files/64/6/Chodacki.pdf.
7. Clark R., Deininger R.A., Protecting the nation's critical infrastructure: the vulnerability of U.S. water supply systems, Journal of Contingencies and Crisis Management 2000, 8 (2): 73–80, doi: 10.1111/1468-5973.00126.
8. Council Directive 2008/114/EC of 8 December 2008 on the identification and designation of European critical infrastructures and the assessment of the reed to improve their protection, Official Journal of the European Union JOL_2008_345_R_0075_01.
9. Cubillo F., Pérez P., Water distribution system risk assessment method, Procedia Engineering 2014, 89: 355-362, https://doi.org/10.1016/j.proeng.2014.11.199.
10. Dubiński J., Mutke G., Tatara T., Muszyński L., Barański A., Kowal T., Zasady stosowania zweryfikowanej górniczej skali intensywności drgań GSIGZWKW-2012 do prognozy i oceny skutków oddziaływania wstrząsów indukowanych eksploatacją złóż węgla kamiennego w zakładach górniczych Kompanii Węglowej S. A. na obiekty budowlane i na ludzi, 2013 – instrukcja.
11. Dulińska J., Oddziaływanie drgań powierzchniowych wywołanych wstrząsami górniczymi w rejonie GZW I LGOM na konstrukcję gazociągu, Wstrząsy górnicze – charakterystyka parametrów drgań oraz kryteria oceny wpływu na obiekty budowlane, Główny Instytut Górnictwa, Katowice, 2010.
12. Dulinska J.M., Jasinska D., Performance of Steel Pipeline with Concrete Coating (Modeled with Concrete Damage Plasticity) under Seismic Wave Passage, Applied Mechanics and Materials 2014, 459: 608-613, doi:10.4028/www.scientific.net/AMM.459.608.
13. EN 1998 Eurocode 8: Design of structures for earthquake resistance.
14. Esposito S., Iervolino I.; PGA and PGV Spatial Correlation Models Based on European Multievent Datasets, Bulletin of the Seismological Society of America 2011, 101 (5): 2532-2541, https://doi.org/10.1785/0120110117.
15. Farmania R., Kakoudakisb K., Behzadianc K., Butler D., Pipe failure prediction in water distribution systems considering static and dynamic factors; Procedia Engineering 2017, 186: 117-126, https://doi.org/10.1016/j.proeng.2017.03.217.
16. Frej A., Zuberek W. M., Local effects in peak accelerations caused by mining tremors in bytom syncline region (Upper Silesia), Acta Geodynamica et Geomaterialia 2008, 5 (2): 115–122, https://www.irsm.cas.cz/materialy/acta_content/2008_02/3_Frej.pdf.
17. Gangl G., Fuchs-Hanusch D., Stadlober E., Kauch P.; Analysis of the failure behaviour of drinking water pipelines; Water Science and Technology: Water Supply 2007, 7 (5-6): 219-225, doi: 10.2166/wst.2011.507.
18. Golik, A. Mendecki, M., Ground-motion prediction equations for induced seismicity in the main anticline and main syncline, Upper Silesian Coal Basin, Poland 2012, 60 (2): 410 – 425, https://doi.org/10.2478/s11600-011-0070-9.

19. Hamdala .F. K., Sagar G. Y., Statistical analysis of pipe breaks in water distribution systems in Ethiopia, the case of Hawassa, IOSR Journal of Mathematics 2016; 12 (3): 127 – 136, http://www.iosrjournals.org/iosr-jm/papers/Vol12-issue3/Version-4/N120304127136.pdf.
20. Hotloś H., Ilościowa ocena wpływu wybranych czynników na parametry i koszty eksploatacji sieci wodociągowych, Wrocław, 2007.
21. Hotloś H., Mielcarzewicz E., Metody oceny udziału szkód górniczych w uszkadzalności sieci wodociągowych, Materiały konferencyjne Rola GPW w systemie zaopatrzenia w wodę dziś i jutro, Górnośląskie Przedsiębiorstwo Wodociągów w Katowicach i PZiTS o/Katowice, Ustroń, 1997.
22. Hotloś H., Mielcarzewicz E., Warunki i ocena niezawodności działania sieci wodociągowych i kanalizacyjnych na terenach górniczych, Monografia, Prace Naukowe Instytutu Inżynierii Ochrony Środowiska Politechniki Wrocławskiej, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 2011.
23. Isoyama R., Ishida E., Yune K., Shirozu R, Seismic damage estimation procedure for water supply pipeline, Proceedings of the twelfth world conference on earthquake engineering 2000, 1762, http://www.iitk.ac.in/nicee/wcee/article/1762.pdf.
24. Joachim K., Kalisz P., 2010: Awarie sieci gazowych na terenach górniczych. Główny Instytut Górnictwa „Górnictwo i Środowisko”2010, 4(1): 95- 105.
25. Kalisz P., Stec K., 2016: Oddziaływanie wstrząsów górniczych na gazociągi, Przegląd Górniczy 2016, 72(10): 1 – 8, http://www.sitg.pl/przegladgorniczy/spis-wydawniczy.html
26. Kurzeja J., Estymacja czasu trwania drgań gruntu generowanych silnymi wstrząsami w kopalniach GZW, Przegląd Górniczy 2016, 7 (1124): 51 – 56.
27. Kurzeja J., Seismometric monitoring in the area of the Piekary Śląskie junction of the A1 motorway in terms of recording the vibrations resulting from mining tremors, Journal of Sustainable Mining 2017, 16: 14 – 23, https://doi.org/10.1016/j.jsm.2017.06.002.
28. Kuś K. i inni, Podstawy projektowania układów i obiektów wodociągowych, Politechnika Śląska, Skrypt uczelniany nr 1854, Gliwice, 1995.
29. Kwietniewski M., Miszta-Kruk K., Piotrowska A., 2011: Wpływ temperatury wody w sieci wodociągowej na jej awaryjność w świetle eksploatacyjnych badań niezawodności, Wydawnictwo Politechniki Krakowskiej, Czasopismo Techniczne 1 – Ś/2011, 1 (108): 113-129.
30. Kwietniewski M., Rak J. Niezawodność infrastruktury wodociągowej i kanalizacyjnej w Polsce. Monografie Komitetu Inżynierii Lądowej i Wodnej PAN, Studia z Zakresu Inżynierii, 67, Warszawa 2010,
31. Kwietniewski M., Roman M., Kłoss-Trębaczkiewicz H.: Niezawodność wodociągów i kanalizacji. Arkady. Warszawa 1993
32. Lasocki, S. (2013), Site specific prediction equations for peak acceleration of ground motion due to earthquakes induced by underground mining in Legnica-Głogów Copper District in Poland, Acta Geophysica 2013, 61(5): 1130-1155, https://doi.org/10.2478/s11600-013-0139-8.
33. Lee D. H., Kim B. H., Lee H., Kong, J. S., 2009: Seismic behavior of a buried gas pipeline under earthquake excitations, Engineering Structures 2009, 31: 1011–1023, https://doi.org/10.1016/j.engstruct.2008.12.012
34. Mahmoodian M., Aryai V., Structural failure assessment of buried steel water pipes subject to corrosive environment, Urban Water Journal 2017 14 (10): 1023 – 1030, http://dx.doi.org/10.1080/1573062X.2017.1325500.

35. Mora-Rodríguez J., Delgado - Galván X., Ramos H. M., López-Jiménez P. A., An overview of leaks and intrusion for different pipe materials and failures, Urban Water Journal 2014, 11 (1): 1-10, http://dx.doi.org/10.1080/1573062X.2012.739630.
36. O'Rourke, M. J., Ayala, G., Pipeline damage due to wave propagation, Journal Geotechnical Engineering 1993, 119: 1490–1498, https://doi.org/10.1061/(ASCE)0733-9410(1993)119:9(1490)
37. O'Rourke M. J., Liu X., Response of buried pipelines subject to earthquake effects, Monograph no. 3, Multidisciplinary Center for Earthquake Engineering Research 1999, 33–57.
38. O'Rourke T. D., Jung J. K., Argyrou C., Underground pipeline response to earthquake-induced ground deformation, Soil Dynamics and Earthquake Engineering 2016, 91: 272 – 283, https://doi.org/10.1016/j.soildyn.2016.09.008
39. O'Rourke T. D., Toprak S., Sano Y., Factors Affecting Water Supply Damage Caused by the Northridge Earthquake, Proceedings of 6th US National Conference on Earthquake Engineering 1998: 1–12.
40. Pietrucha-Urbanik K, Studziński A., Case study of failure simulation of pipelines conducted in chosen water supply system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19 (3): 317–323, http://dx.doi.org/10.17531/ein.2017.3.1.
41. Pilch R., Szybka J., Tuszyńska A., Application of factoring and time-space simulation methods for assessment of the reliability of water-pipe networks, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (2): 253-258, http://www.ein.org.pl/sites/default/files/2014-02-12.pdf.
42. Pineda-Porras O., Ordaz M., A new seismic intensity parameter to estimate damage in buried pipeline due to seismic wave propagation, Journal of Earthquake Engineering 2007, 11: 773–786, http://dx.doi.org/10.1080/13632460701242781.
43. Rak J., Bezpieczeństwo systemów zaopatrzenia w wodę, Polska Akademia Nauk, Instytut Badań Systemowych, Warszawa 2009.
44. Rak J., Wybrane aspekty bezpieczeństwa systemów wodociągowych, Oficyna Wydawnicza Politechniki Rzeszowskie 2015.
45. Rak J., Tchurzewska-Cieślak B., Studziński A., Pietrucha-Urbanik K., Boryczko K., Niezawodność i bezpieczeństwo systemów zbiorowego zaopatrzenia w wodę, Oficyna Wydawnicza Politechniki Rzeszowskiej, Rzeszów 2012.
46. Rezaei H., Ryan B., Stoianov I.; Pipe failure analysis and impact of dynamic hydraulic conditions in water supply networks; Procedia Engineering 2015, 119: 253 – 262, https://doi.org/10.1016/j.proeng.2015.08.883.
47. Scheidegger A, Leitão J P, Scholten L., Statistical failure models for water distribution pipes – A review from a unified perspective, Water Research 2015, 83: 237–247, https://doi.org/10.1016/j.watres.2015.06.027.
48. Si H., Midorikava S., New attenuation relations for peak ground acceleration and velocity considering effects of faulty type and site condition, Journal of structural and construction engineering 1999, 64 (523): 63 – 70, http://doi.org/10.3130/aijs.64.63_2.
49. Takada S. and Tanabe K., 1987: Three-dimensional seismic response analysis of buried continuous or jointed pipelines, Journal of Pressure Vessel Technology 1987, 109: 80–87, doi:10.1115/1.3264859.
50. Ustawa z dnia 26 kwietnia 2007 r. o zarządzaniu kryzysowym. Dz. U. 2007 Nr 89 poz. 590 wraz z póź. zm.

51. Ustawa z dnia 7 czerwca 2001 r. o zbiorowym zaopatrzeniu w wodę i zbiorowym odprowadzaniu ścieków Dz. U. 2017 poz. 328 wraz z póź. zm.
52. Wang, L. R.-L., Cheng, K.-M., Seismic response behavior of buried pipelines, Journal of Pressure Vessel Technology 1979, 101: 21–30, doi:10.1115/1.3454594.
53. Wieczysty A.: Niezawodność systemów wodociągowych i kanalizacyjnych Cz. I i II, Teoria niezawodności i jej zastosowania, Wydawnictwo Politechniki Krakowskiej, Kraków 1990.
54. Zasada działania rejestratora drgań AMAX-GSI, Instrukcja obsługi, Dokumentacja techniczno - ruchowa, Labolatorium Sejsmologii i sejsmiki górniczej.
55. Zembaty Z., Rockburst induced ground motion – a comparative study, Soil Dynamics and Earthquake Engineering 2004, 24 (1): 11 – 23, https://doi.org/10.1016/j.soildyn.2003.10.001.