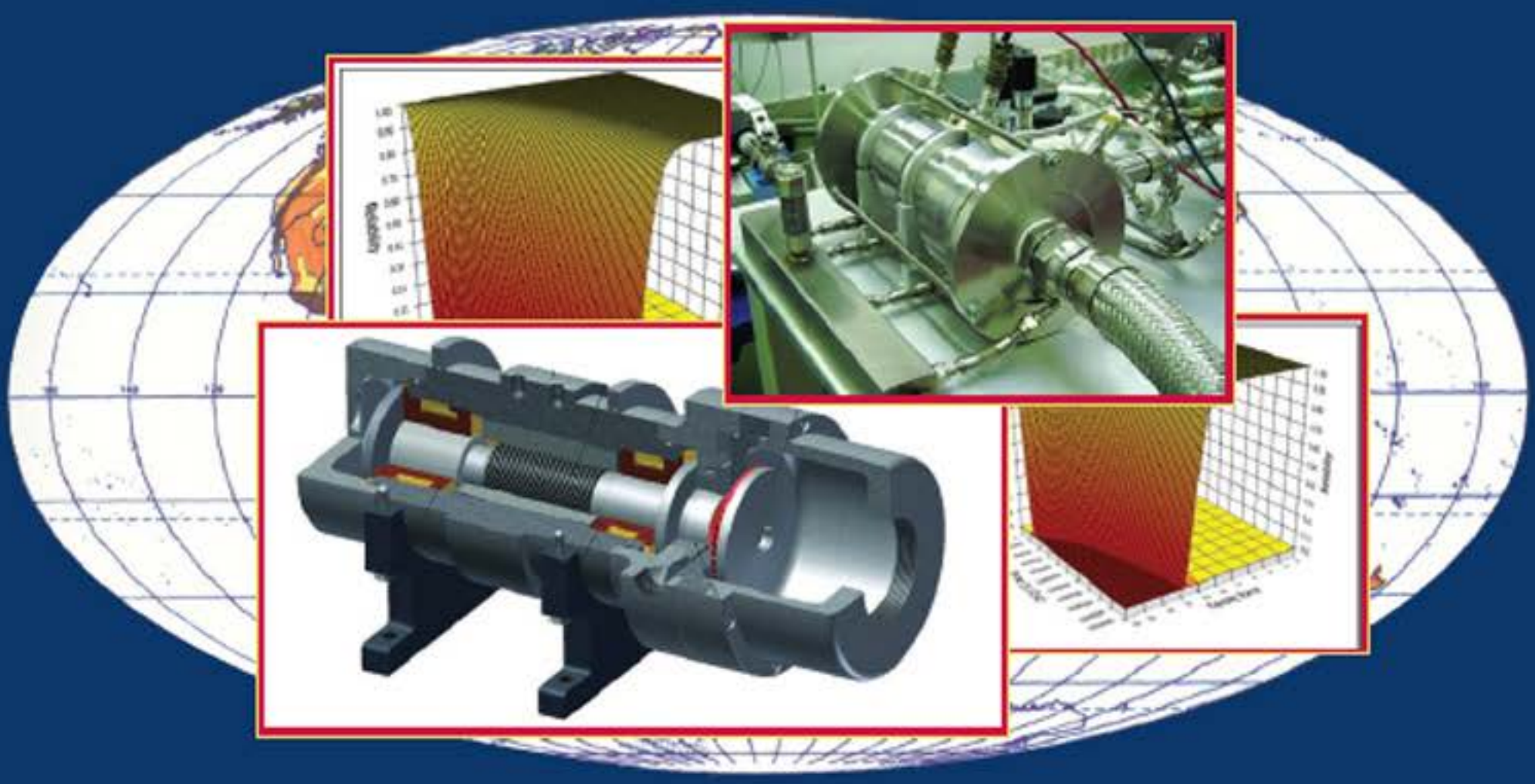
Vol. 21. No 4, 2019

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

HAN T, JIANG D. **Fault diagnosis of multistage centrifugal pump unit using non-local means-based vibration signal denoising**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 539–545, http://dx.doi.org/10.17531/ein.2019.4.1.

In real industry environment, the signal characteristics of multistage centrifugal pump vibration signal are easily submerged by strong background noise. To settle this problem, the nonlocal means (NLM) approach is proposed for the denoising of multistage centrifugal pump in this paper. Utilizing the similarity theory, the NLM method has achieved a wide range of applications in the fields of image processing and biomedical signal denoising. Due to the periodic characteristics and redundancy, NLM is successfully applied to the de-noising of 1-D machinery vibration signal. The numerical simulation experiments with different SNRs verify the effectiveness and the superiority of the proposed method. Besides, the selection principles of core parameters in NLM are discussed. The real engineering cases analysis demonstrates that the NLM can effectively filter out the background noise and realize the weak fault feature enhancements. The proposed noise reduction method is superior to traditional wavelet coefficient method.

HAN T, JIANG D. **Diagnozowanie uszkodzeń wielostopniowej pompy odśrodkowej z wykorzystaniem metody odszumiania sygnału drgań w oparciu o średnie nielokalne**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 539–545, http://dx.doi.org/10.17531/ein.2019.4.1.

W rzeczywistym środowisku przemysłowym, charakterystyki sygnału drgań wielostopniowej pompy odśrodkowej są zagłuszane przez silny szum tła. Problem ten można rozwiązać stosując zaproponowane w niniejszej pracy podejście oparte na algorytmie średnich nielokalnych (non-local means, NLM). Wykorzystująca teorię podobieństwa metoda NLM znajduje szeroki zakres zastosowań w dziedzinie przetwarzania obrazu i odszumiania sygnałów biomedycznych. Dzięki okresowemu charakterowi i redundancji sygnałów, NLM można z powodzeniem stosować do usuwania szumu jednowymiarowego sygnału drgań maszyn. Skuteczność proponowanej metody i jej przewagę nad stosowanymi dotychczas rozwiązaniami zweryfikowano na podstawie eksperymentów symulacyjnych z uwzględnieniem różnych stosunków sygnału do szumu (SNR). Ponadto omówiono zasady wyboru podstawowych parametrów NLM. Analiza przypadków inżynierskich pokazuje, że NLM pozwala skutecznie odfiltrowywać szumy tła i wzmacniać słabe symptomy akustyczne uszkodzenia. Proponowana metoda redukcji szumów przewyższa tradycyjną metodę współczynnika falkowego.

CHERNETS M. **Method of calculation of tribotechnical characteristics of the metal-polymer gear, reinforced with glass fiber, taking into account the correction of tooth**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 546–552, http://dx.doi.org/10.17531/ein.2019.4.2.

The paper proposes a new method for calculating the service life, wear and contact pressures of metal-polymer gear drives with a correction profile. The effects of height and angular modification in a gear drive made of dispersive glass fibre-reinforced polyamide and steel on its contact and tribocontact parameters are determined. A numerical solution obtained for the gear with height correction has shown that the life of such gear is the longest when the profile correction coefficients $x_1 = -x_2 = 0.1$. It has been found that the service life of the gear with angular correction is shorter than that of the gear with correction height. The effects of gear tooth height and angular correction on maximum contact pressures and pinion wear are examined and determined.

CHERNETS M. **Metoda obliczeniowa tribotechnicznych charakterystyk przekładni zębatych metal-polimerowych z poliamidu wzmocnionego włóknem szklanym z uwzględnieniem korekcji zębów**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 546–552, http://dx.doi.org/10.17531/ein.2019.4.2.

Predstawiono opracowaną nową metodę obliczeniową resursu, zużycia oraz nacisków stykowych przekładni walcowej metal – polimerowej z korekcja uzębienia. Dla przekładni z kołami zębatymi z poliamidu wzmocnionego dyspersyjnym włóknem szklanym i stali zostało przeprowadzone oszacowanie wpływu korekcji technologicznej oraz konstrukcyjnej uzębienia na wskazane parametry kontaktu oraz tribokontaktu. Na podstawie numerycznego rozwiązania zagadnienia dla przypadku korekcji technologicznej zębów kół określono, że największa trwałość przekładni będzie, gdy współczynniki korekcji $x_1 = -x_2 = 0.1$. Ustalono, że wtedy przy korekcji konstrukcyjnej zębów trwałość przekładni będzie mniejsza nieżeli przy korekcji technologicznej. Został przebadany charakter wpływu korekcji technologicznej oraz konstrukcyjnej zębów na maksymalne naciski stykowe, zużycie zębnika oraz ustalono jego prawidłowości.

PAŚ J, KLIMCZAK T. **Selected issues of the reliability and operational assessment of a fire alarm system**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 553–561, http://dx.doi.org/10.17531/ein.2019.4.3.

The article discusses an analysis of the operational and reliability issues, which regards selected fire alarm systems (FAS) exhibiting different functional structures. These systems are operated in a vast transport area, within a specific environment. We can distinguish three basic structures of these systems – focused, dispersed and mixed. A given system functional structure, utilized within a facility (a given area) is a function depending on the configuration, internal connections of elements and devices, and a developed fire scenario. The application of a given system structure for fire protection also depends on the legislation determining the approval of a facility (area) for use. The process of executing a scenario in the event of a fire is ensured by an algorithm implemented in the alarm central unit and other elements of the system. The implementation of all the system requirements specified within a given procedure algorithm depends on, e.g., an appropriate reliability structure and environmental conditions. The article analyses the operational process of selected FAS, which are operated within a vast transport area. It discusses the actual results of the operational process tests, e.g., repair and damage durations. Next, operational relationship graph, taking into account the conducted operational test, was developed. This enabled the determination of relationships that allow to specify the operating and reliability parameters in terms of a FAS staying in the states distinguished for the research. The FAS test methodology presented in the article, owing to meeting specific performance requirements, can be used in the course of developing a fire scenario and designing systems, taking into account various available technical solutions.

PAŚ J, KLIMCZAK T.**Wybrane zagadnienia oceny niezawodnościowo-eksploatacyjnej systemów sygnalizacji pożaru**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 553–561, http://dx.doi.org/10.17531/ein.2019.4.3.

W artykule przeprowadzono analizę problemów eksploatacyjnych i niezawodnościowych, która dotyczy wybranych systemów sygnalizacji pożaru (SSP) o różnej strukturze funkcjonalnej. Systemy te są użytkowane na rozległym obszarze transportowym, w określonym środowisku. Można wyróżnić trzy podstawowe struktury tych systemów - skupiona, rozproszona i mieszana. Dany rodzaj struktury funkcjonalnej systemu, który jest użytkowany w obiekcie (na danym obszarze) jest funkcją zależną od konfiguracji, wewnętrznych połączeń elementów i urządzeń oraz opracowanego scenariusza postępowania na wypadek pożaru. Zastosowanie danej struktury systemu do ochrony pożarowej zależy także od przepisów prawnych warunkujących dopuszczenie danego obiektu (obszaru) do użytkowania. Proces realizacji scenariusza w czasie pożaru jest gwarantowany przez algorytm zaimplementowany w centrali alarmowej oraz innych elementach systemu. Realizacja wszystkich wymagań wobec systemu określonych w danym algorytmie postępowania uwarunkowana jest np. odpowiednią strukturą niezawodnościową i warunkami środowiskowymi. W artykule przedstawiono analizę procesu eksploatacji wybranych SSP, które są użytkowane na obszarze transportowym. Zaprezentowano rzeczywiste wyniki badań procesu eksploatacji, np. czasy trwania naprawy oraz uszkodzenia. Następnie opracowano graf relacji eksploatacyjnych z uwzględnieniem przeprowadzonych badań eksploatacyjnych. Umożliwiło to wyznaczenie zależności pozwalających na określenie parametrów eksploatacyjnych i niezawodnościowych przebywania SSP w wyróżnionych do rozważań stanach. Przedstawiona w artykule metodyka badania SSP ze względu na spełnienie określonych wymagań eksploatacyjnych może być użyta podczas opracowywania scenariusza pożarowego oraz projektowania systemów z uwzględnieniem różnych dostępnych rozwiązań technicznych.

KUCHARSKA B. **Identification of surface stress in the exhaust system pipe made by hydroforming technology based on diffractometric measurements**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 562–566, http://dx.doi.org/10.17531/ein.2019.4.4.

In the work identificaton of surface stresses in the exhaust pipe made of Cr-Ni steel shaped with hydroforming technology. Stresses were determined by the non-destructive x-ray method $\sin^2\psi$. A complex state of tensile stresses with values in the range of 69÷240 MPa for circumferential stresses and 26÷290 MPa for longitudinal stresses was found on the surface of the pipe. The distribution of stresses on the circumference and length of the pipe was analyzed on the basis of coefficients of variation and wall thickness. A relationship was found between the value of surface stress and the wall thickness of the pipe. The highest stresses occurred in the areas of the pipe where

KUCHARSKA B. **Identyfikacja naprężeń powierzchniowych w rurze do układu wydechowego wykonanej technologią hydroformowania na podstawie pomiarów dyfraktometrycznych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 562–566, http://dx.doi.org/10.17531/ein.2019.4.4.

W pracy dokonano identyfikacji naprężeń powierzchniowych w rurze wydechowej ze stali Cr-Ni kształtowanej technologią hydroformowania. Naprężenia wyznaczono nieniszczącą rentgenowską metodą $\sin^2\psi$. Na powierzchni rury stwierdzono złożony stan naprężeń rozciągających o wartościach z zakresu 69÷240 MPa dla naprężeń obwodowych i 26÷290 MPa dla naprężeń wzdłużnych. Rozłożenie naprężeń na obwodzie i długości rury analizowano na podstawie współczynników zmienności i grubości ścianki. Stwierdzono zależność pomiędzy wartością naprężeń powierzchniowych a grubością ścianki rury. Największe naprężenia występowały w obszarach rury gdzie grubość ścianki była najsilniej

the thickness of the wall was reduced the most. In the central part of the pipe, where the wall thickness reduction was the smallest, the stresses were also the smallest, but they were characterized by the highest dispersion of value. The distribution of surface stresses determined by diffractometric method was compared with the model of deformation of the pipe generated numerically.

zredukowana. W centralnej części rury gdzie redukcja grubości ścianki była najmniejsza naprężenia również były najmniejsze, ale cechowały się największym rozproszeniem wartości. Rozłożenie naprężeń powierzchniowych wyznaczonych metodą dyfraktometryczną porównano z modelem odkształceń w rurze wygenerowanym numerycznie.

KANG R, WANG J, CHENG J, CHEN J, PANG Y. **Intelligent forecasting of automatic train protection system failure rate in China high-speed railway**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 567–576, http://dx.doi.org/10.17531/ein.2019.4.5.

Intelligent and personalized dynamic maintenance and spare parts configuration of high-speed railway have been the main trend to guarantee the safety capability of trains. In this paper, a new Automatic Train Protection (ATP) system failure rate calculation method is proposed, and the delay time and embedded dimension are determined by C-C algorithm. Then the phase space is reconstructed from one-dimensional time series to high-dimensional space. Based on chaotic characteristics of failure rate, a short-term intelligent forecasting model of failure rate of ATP system is established. The actual failure statistics from 2010 to 2018 are used as samples to train and test the validity of the model. From prediction results, it shows that the proposed chaos prediction model has an accuracy of 99.71%, which is better than the support vector machine model. Through the intelligent prediction of failure rate, this paper solves the maintenance inflexibility and imbalance of supply and demand of spare parts configuration.

KANG R, WANG J, CHENG J, CHEN J, PANG Y. **Inteligentne prognozowanie intensywności uszkodzeńautomatycznego systemu ochrony pociągów kolei dużych prędkości w Chinach**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 567–576, http://dx.doi.org/10.17531/ein.2019.4.5.

Inteligentna i spersonalizowana dynamiczna konserwacja i konfiguracja części zamiennych pociągów kolei dużych prędkości stanowią ostatnio główny trend w zakresie zapewniania bezpieczeństwa pociągów. W niniejszym artykule zaproponowano nową metodę obliczania intensywności uszkodzeń systemu Automatycznej Ochrony Pociągu (ATP), a czas opóźnienia i wymiar zanurzenia określano za pomocą algorytmu CC. Następnie, przestrzeń fazową przekształcono z jednowymiarowego szeregu czasowego do przestrzeni wielowymiarowej. Opierając się na chaotycznych charakterystykach intensywności uszkodzeń, utworzono model krótkoterminowego inteligentnego prognozowania awaryjności systemu ATP. Do uczenia modelu i weryfikacji jego trafności wykorzystano rzeczywiste dane statystyczne dotyczące awarii pociągów z lat 2010–2018. Z wyników prognoz wynika, że proponowany model predykcji, oparty na teorii chaosu, cechuje się dokładnością na poziomie 99,71%, czyli wyższą niż model maszyny wektorów nośnych. Dając możliwość inteligentnej predykcji intensywności uszkodzeń, niniejsza praca rozwiązuje problem braku elastyczności w utrzymaniu ruchu pociągów oraz braku równowagi między podażą a popytem na części zamienne.

JAMROZIAK K, KWASNIOWSKI S, KOSOBUDZKI M, ZAJAC P. **Analysis of heat exchange in the powertrain of a road vehicle with a retarder**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 577–584, http://dx.doi.org/10.17531/ein.2019.4.6.

The paper presents a heat exchange model for the cooling system of any complex, physical system. Verification of the correctness of the theoretical model was carried out on the example of a vehicle with a combustion engine and additionally equipped with a hydraulic retarder. The results of laboratory tests, which were carried out on an engine test bench, were also performed for the above mentioned powertrain, so as to compare the results of modelling with the results of the tests. Determining the operating parameters of the components of the cooling system aimed at protecting the entire powertrain against overheating is a key task. Theoretical analysis of heat exchange in the powertrain of a road vehicle was carried out, with particular emphasis on the hydraulic retarder (a device braking the vehicle during a descent on roads with a high gradient of the road, mandatory according to the ADR convention). The subject of the study was a mathematical model of a complex cooling system developed by the authors, described by means of balance equations and differential equations. This model was tested with the use of the Matlab-Simulink suite for given load parameters of the cooling system, which were used in tests on an engine test bench. The values of coefficients describing the thermal state of the powertrain were obtained. Simulations were performed for different variants of technical parameters of the expanded cooling system. In this way, individual units and components of the cooling system were optimized so that it fulfilled its role in the assumed operating conditions and the ecologization of emission of energy sources (fuel) and harmful substances.

JAMROZIAK K, KWASNIOWSKI S, KOSOBUDZKI M, ZAJAC P. **Analiza wymiany ciepła w układzie napędowym pojazdu drogowego z retarderem**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 577–584, http://dx.doi.org/10.17531/ein.2019.4.6.

Praca zawiera model wymiany ciepła układu chłodzenia dla złożonego, dowolnego układu fizycznego. Weryfikację poprawności modelu teoretycznego przeprowadzono na przykładzie pojazdu z silnikiem spalinowym oraz dodatkowo wyposażonym w retarder hydrauliczny. Wyniki badań laboratoryjnych, które przeprowadzono na hamowni również wykonano dla w/w zespołu napędowego, tak aby porównać wyniki modelowania z wynikami badań. Ustalenie parametrów eksploatacji elementów układu chodzenia, którego celem jest zabezpieczenie całego układu napędowego przed przegrzaniem to kluczowe zadanie. Przeprowadzono teoretyczną analizę wymiany ciepła w układzie napędowym pojazdu drogowego ze szczególnym uwzględnieniem zwalniacza hydraulicznego (urządzenia hamującego pojazd podczas zjazdu na drogach o dużym pochyleniu jezdni, obowiązkowe wg konwencji ADR). Przedmiotem badań był opracowany przez autorów model matematyczny rozbudowanego układu chłodzenia opisany za pomocą równań bilansowych i równań różniczkowych. Model ten testowano z wykorzystaniem pakietu Matlab-Simulink dla zadanych parametrów obciążenia układu chłodzenia, które wykorzystywano w badaniach w stanowiskowych na hamowni silnikowej. Uzyskano wartości współczynników opisujących stan cieplny jednostki napędowej. Symulacje wykonano dla różnych wariantów parametrów technicznych rozbudowanego układu chłodzenia. W ten sposób optymalizowano poszczególne zespoły i podzespoły układu chłodzenia, tak aby spełniał on swoją rolę w zakładanych warunkach eksploatacji i ekologizację emisji źródeł energii (paliwa) i szkodliwych substancji.

CASTILLA-GUTIÉRREZ J, FORTES JC, PULIDO-CALVO I. **Analysis, evaluation and monitoring of the characteristic frequencies of pneumatic drive unit and its bearing through their corresponding frequency spectra and spectral density**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 585–591, http://dx.doi.org/10.17531/ein.2019.4.7.

This article shows the results of the study of the characteristic frequencies of pneumatic drive equipment and its suspension bearing. The analysis approaches one of the most important requirements of the industrial sector, which seeks to be recognised by the efficiency and performance of its equipment when compared to its coming economic competitors. For data collection and we have followed the ISO 10816 standards, thus using the values of speed in RMS, aiming to reduce the masking of these signals that occurs depending on whether they are high or low frequencies. The study will respond to one of the most important requirements found in the predictive and preventive control of industrial sites. The problem of the predictive systems of maintenance of equipment with bearings lies in the number of monitoring and analysis points that generate a high cost in time and human resources. The aim will be to determine which of all the study frequencies is the most significant and in which position and measurement axis has the biggest impact. To do this, we will analyse the rotation frequency of the blowing machine, the resulting frequency of all the frequencies, the frequency of the impulsion blades and finally the frequency of the bearing. The study would be able to predict when our equipment is going to suffer a failure, reducing the control points and the cost.

CASTILLA-GUTIÉRREZ J, FORTES JC, PULIDO-CALVO I. **Analiza, ocena i monitorowanie częstotliwości charakterystycznych pneumatycznego zespołu napędowego i wchodzącego w jego skład łożyska na podstawie widm częstotliwości oraz gęstości widmowej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 585–591, http://dx.doi.org/10.17531/ein.2019.4.7.

W artykule przedstawiono wyniki badań częstotliwości charakterystycznych napędu pneumatycznego i wchodzącego w jego skład łożyska zawieszenia. Analiza przybliża jedno z najważniejszych wymagań sektora przemysłowego, w którym dąży się do tego by wyróżniać się na tle konkurencji sprawnością i wydajnością urządzeń. Przy zbieraniu danych postępowaliśmy zgodnie ze normą ISO 10816, wykorzystując średnie prędkości kwadratowe, co pozwoliło zmniejszyć maskowanie sygnałów, które występuje w zależności od tego, czy mamy do czynienia z wysokimi czy niskimi częstotliwościami. Badanie stanowi odpowiedź na jeden z najważniejszych wymogów w zakresie kontroli predykcyjnej i prewencyjnej obiektów przemysłowych. Problemem systemów konserwacji predykcyjnej sprzętu, w którego skład wchodzą łożyska jest duża ilość punktów kontrolnych, które generują wysokie koszty jeśli chodzi o czas i zasoby ludzkie. Celem pracy było określenie, które ze wszystkich badanych częstotliwości są najistotniejsze oraz dla których częstotliwości pozycja i oś pomiaru mają największe znaczenie. W tym celu przeanalizowano częstotliwość obrotową analizowanej dmuchawy, częstotliwość wynikową wszystkich częstotliwości, częstotliwość łopatek oraz częstotliwość łożyska.

KOSOBUDZKI M, STAŃCO M. **Problems in assessing the durability of a selected vehicle component based on the accelerated proving ground test**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 592–598, http://dx.doi.org/10.17531/ein.2019.4.8.

The paper presents the results of the analysis of the durability of elastic elements occurring in the special-purpose 4x4 off-road truck suspension using data obtained during an accelerated proving ground test conducted during off-road driving. The limitations in access to material data present at the stage of the initial selection of the component (lack of fatigue strength data) are indicated and an alternative analytical method for fatigue strength estimation is given. The differences in the obtained results and their most important sources are pointed out. A method for using a generalized durability index d as a parameter independent of the subassembly material data is also described. The indicator can be used to assess the influence of resultant loads (recorded) appearing during the vehicle operation in the determined road conditions on the durability of the subassembly under study and to associate their value with the type of the test road section.

KOSOBUDZKI M, STAŃCO M. **Problemy oceny trwałości wybranego elementu pojazdu na podstawie przyspieszonego testu przebiegowego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 592–598, http://dx.doi.org/10.17531/ein.2019.4.8.

W artykule przedstawiono wyniki analizy trwałości elementów sprężystych występujących w zawieszeniu specjalnego terenowego pojazdu ciężarowego 4x4 wykorzystując dane uzyskane podczas przyspieszonego testu drogowego przeprowadzonego podczas jazdy off-road. Wskazano na występujące ograniczenia w dostępie do danych materiałowych jakie są obecne na etapie wstępnego doboru podzespołu (brak danych wytrzymałości zmęczeniowej) oraz podano alternatywną analityczną metodę szacowania wytrzymałości zmęczeniowej. Wskazano na powstające różnice w uzyskanych wynikach oraz na najważniejsze ich źródła. Przedstawiono również sposób wykorzystania uogólnionego wskaźnika trwałości d jako parametru niezależnego od danych materiałowych podzespołu, który można wykorzystać do oceny wpływu obciążeń wynikowych (rejestrowanych) powstających podczas ruchu pojazdu w ustalonych warunkach drogowych na trwałość analizowanego podzespołu i powiązać ich wartość z rodzajem testowego odcinka drogowego.

YOU L, ZHANG J, Li Q, Ye N. **Structural reliability analysis based on fuzzy random uncertainty**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 599–609, http://dx.doi.org/10.17531/ein.2019.4.9.

To address the fuzzy random uncertainty in structural reliability analysis, a novel method for obtaining the membership function of fuzzy reliability is proposed on the two orders four central moments (TOFM) method based on envelope distribution. At each cut level, the envelope distribution is first constructed, which is a new expression to describe the bound of the fuzzy random variable distribution. The central moments of the bound distribution are determined by generating samples from the envelope distribution, and they are used to calculate the central moments of the limit state function based on the first two orders of the Taylor expansion. Thereafter, the modern approximation method is used to approximate the polynomial expression for the limit state function probability density function (PDF) by considering the central moments as constraint conditions. Thus, the reliability boundaries can be calculated under the considered cut level, and the membership function of the fuzzy reliability is subsequently obtained. Three examples are evaluated to demonstrate the efficiency and accuracy of the proposed method. Moreover, a comparison is made between the proposed method, Monte Carlo simulation (MCS) method, and fuzzy first-order reliability method (FFORM). The results show the superiority of the proposed method, which is feasible for the analysis of structural reliability with fuzzy randomness.

YOU L, ZHANG J, Li Q, Ye N. **Analiza niezawodności strukturalnej w oparciu o rozmytą niepewność losową**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 599–609, http://dx.doi.org/10.17531/ein.2019.4.9.

W pracy przedstawiono metodę, która pozwala na uwzględnienie rozmytej niepewności losowej w strukturalnej analizie niezawodności. Zaproponowana metoda określania funkcji przynależności niezawodności rozmytej wykorzystuje cztery momenty centralne dwóch rzędów czy czwarte momenty centralne drugiego rzędu obliczane w oparciu o rozkład obwiedni. Dla każdego poziomu cięcia, najpierw konstruuje się rozkład prawdopodobieństwa obwiedni, za pomocą którego opisuje się granice rozkładu rozmytych zmiennych losowych, a momenty centralne rozkładu ograniczonego wyznacza się poprzez generowanie prób z rozkładu obwiedni. Następnie, stosując nowoczesną metodę optymalnej aproksymacji, otrzymuje się aproksymowane wyrażenie wielomianowe funkcji gęstości prawdopodobieństwa rozkładu obwiedni, gdzie momenty centralne stanowią warunki ograniczające, które pozwalają aproksymować niezawodność za pomocą rozwinięcia Taylora drugiego rzędu funkcji stanu granicznego. W ten sposób granice niezawodności oblicza się na rozważanym poziomie cięcia, a następnie otrzymuje się funkcję przynależności niezawodności rozmytej. W artykule przeanalizowano trzy przykłady, na podstawie których wykazano skuteczność i trafność proponowanej metody. Przeprowadzono także porównanie z metodą symulacji Monte Carlo oraz metodą analizy rozmytej niezawodności pierwszego rzędu. Wyniki wskazują na wyższość omawianej metody, która pozwala analizować niezawodność strukturalną w warunkach losowości rozmytej.

MAIOR CBS, CHAGAS MOURA M, LINS ID. **Particle swarm-optimized support vector machines and pre-processing techniques for remaining useful life estimation of bearings**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 610–618, http://dx.doi.org/10.17531/ein.2019.4.10.

The useful life time of equipment is an important variable related to system prognosis, and its accurate estimation leads to several competitive advantage in industry. In this paper, Remaining Useful Lifetime (RUL) prediction is estimated by Particle Swarm optimized Support Vector Machines (PSO+SVM) considering two possible pre-processing techniques to improve input quality: Empirical Mode Decomposition (EMD) and Wavelet Transforms (WT). Here, EMD and WT coupled with SVM are used to predict RUL of bearing from the IEEE PHM Challenge 2012 big dataset. Specifically, two cases were analyzed: considering the complete vibration dataset and considering truncated vibration dataset. Finally, predictions provided from models applying both pre-processing techniques are compared against results obtained from PSO+SVM without any pre-processing approach. As conclusion, EMD+SVM presented more accurate predictions and outperformed the other models.

MAIOR CBS, CHAGAS MOURA M, LINS ID. **Zastosowanie maszyn wektorów nośnych zoptymalizowanych metodą roju cząstek oraz technik przetwarzania wstępnego do oceny pozostałego okresu użytkowania łożysk**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 610–618, http://dx.doi.org/10.17531/ein.2019.4.10.

Okres użytkowania sprzętu jest ważną zmienną związaną z prognozowaniem pracy systemu, a możliwość jego dokładnej oceny daje zakładom przemysłowym znaczną przewagę konkurencyjną. W tym artykule pozostały czas pracy (Remaining Useful Life, RUL) szacowano za pomocą maszyn wektorów nośnych zoptymalizowanych rojem cząstek (SVM+PSO) z uwzględnieniem dwóch technik przetwarzania wstępnego pozwalających na poprawę jakości danych wejściowych: empirycznej dekompozycji sygnału (Empirical Mode Decomposition, EMD) oraz transformat falkowych (Wavelet Transforms, WT). W niniejszej pracy, EMD i falki w połączeniu z SVM wykorzystano do prognozowania RUL łożyska ze zbioru danych IEEE PHM Challenge 2012 Big Dataset. W szczególności, przeanalizowano dwa przypadki: uwzględniający kompletny zestaw danych o drganiach oraz drugi, biorący pod uwagę okrojoną wersję tego zbioru. Prognozy otrzymane na podstawie modeli, w których zastosowano obie techniki przetwarzania wstępnego porównano z wynikami uzyskanymi za pomocą PSO + SVM bez wstępnego przetwarzania danych. Wyniki pokazały, że model EMD + SVM generował dokładniejsze prognozy i tym samym przewyższał pozostałe badane modele.

BABISHIN V, TAGHIPOUR S. **An algorithm for estimating the effect of maintenance on aggregated covariates with application to railway switch point machines**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 619–630, http://dx.doi.org/10.17531/ein.2019.4.11.

We propose an algorithm for estimating the effectiveness of maintenance on both age and health of a system. One of the main contributions is the concept of virtual health of the device. It is assumed that failures follow a nonhomogeneous Poisson process (NHPP) and covariates follow the proportional hazards model (PHM). In particular, the effect of maintenance on device's age is estimated using the Weibull hazard function, while the effect on device's health and covariates associated with condition-based monitoring (CBM) is estimated using the Cox hazard function. We show that the maintenance effect on the health indicator (HI) and the virtual HI can be expressed in terms of the Kalman filter concepts. The HI is calculated from Mahalanobis distance between the current and the baseline condition monitoring data. The effect of maintenance on both age and health is also estimated. The algorithm is applied to the case of railway point machines. Preventive and corrective types of maintenance are modelled as different maintenance

BABISHIN V, TAGHIPOUR S. **Algorytm do oceny wpływu konserwacji na zagregowane zmienne towarzyszącei jego zastosowanie w odniesieniu do kolejowych napędów zwrotnicowych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 619–630, http://dx.doi.org/10.17531/ein.2019.4.11.

W artykule zaproponowano algorytm służący do szacowania skuteczności utrzymania ruchu w odniesieniu do wieku i stanu technicznego (kondycji) systemu. Główny wkład proponowanej metody stanowi koncepcja wirtualnego stanu urządzenia. Metoda zakłada, że uszkodzenia można zamodelować za pomocą niejednorodnego procesu Poissona, a zmienne towarzyszące za pomocą modelu proporcjonalnego hazardu. Mówiąc precyzyjniej, wpływ konserwacji na wiek urządzenia szacuje się z wykorzystaniem funkcji hazardu Weibulla, natomiast wpływ na stan urządzenia i zmienne towarzyszące związane z monitorowaniem stanu ocenia się stosując funkcję hazardu Coxa. W artykule pokazujemy, że wpływ konserwacji na wskaźnik stanu i wskaźnik stanu wirtualnego można wyrazić w kategoriach filtra Kalmana. Wskaźnik stanu oblicza się na podstawie odległości Mahalanobisa między bieżącymi a początkowymi danymi z monitorowania stanu. Ocenia się także wpływ utrzymania na wiek i kondycję systemu. Proponowany algorytm

effect parameters. Using condition monitoring data, the HI is calculated as a scaled Mahalanobis distance. We derive reliability and likelihood functions and find the least squares estimates (LSE) of all relevant parameters, maintenance effect estimates on time and HI, as well as the remaining useful life (RUL).

zastosowano w odniesieniu do napędów zwrotnicowych. Zapobiegawcze i naprawcze typy konserwacji zamodelowano jako różne parametry utrzymania ruchu. Korzystając z danych z monitorowania stanu, obliczono wskaźnik stanu jako skalowaną odległość Mahalanobisa. Wyprowadzono funkcje niezawodności i wiarygodności oraz obliczono metodą najmniejszych kwadratów szacunkowe wielkości wszystkich istotnych parametrów, a także szacunkowy wpływ konserwacji na wskaźniki czasu i stanu technicznego oraz pozostały okres użytkowania (RUL).

ČEPIN M. **Evaluation of the importance factors of the power plants within the power system reliability evaluation**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 631–637, http://dx.doi.org/10.17531/ein.2019.4.12.

The objective of the paper is to develop the reliability importance factors which could identify the plants, which more or less contribute to the increased or decreased power system reliability. One group of importance factors could identify the plants, which with their increased availability would notably increase the power system reliability The other group of the importance factors could identify the plants, which with their reduced availability would notably reduce the power system reliability. The importance factors have been developed. An example of regional power system was considered for the case study. The results identify the power plants which are more susceptible to increase of the loss of load expectation and thus to decreasing of the power system reliability, if their availability is reduced. Similarly, the results identify the power plants which are more susceptible to decreasing of the loss of load expectation and thus to increase of the power system reliability, if their availability is increased. The lists of important factors can serve as a standpoint for inclusion of the power system reliability role within the power system to the planning activities. The lists of importance factors can represent the standpoint for the power system operator to reward the improvement of the power plant availability or to penalize the reduced power plant availability.

ČEPIN M. **Ocena współczynników ważności elektrowni w ramach oceny niezawodności systemu elektroenergetycznego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 631–637, http://dx.doi.org/10.17531/ein.2019.4.12.

Celem artykułu jest opracowanie współczynników ważności niezawodności, pozwalających identyfikować elektrownie, które w mniejszym lub większym stopniu przyczyniają się do zwiększenia lub zmniejszenia niezawodności systemu elektroenergetycznego. Opracowano dwie grupy takich współczynników: jedne – służące do identyfikacji elektrowni, które przy zwiększonej gotowości mogą znacznie zwiększać niezawodność systemu elektroenergetycznego, i drugie – do identyfikacji tych elektrowni, których zmniejszona gotowość może znacznie zmniejszać niezawodność sieci elektroenergetycznej. Jako studium przypadku rozważano przykład regionalnego systemu elektroenergetycznego. Wyniki pozwalają na wskazanie elektrowni, które są bardziej podatne na zwiększenie oczekiwanego czasu deficytu mocy (niepokrycia zapotrzebowania), a tym samym mogą bardziej przyczyniać się do zmniejszenia niezawodności systemu elektroenergetycznego, w przypadku spadku ich gotowości. Podobnie, uzyskane wyniki pozwalają ustalić, dla których elektrowni prawdopodobieństwo zmniejszenia oczekiwanego czasu deficytu mocy jest większe, a tym samym, które elektrownie mogą bardziej przyczyniać się do zwiększenia niezawodności systemu elektroenergetycznego, gdy zwiększy się ich gotowość. Listy współczynników ważności mogą służyć jako punkt odniesienia dla działań planistycznych, a także jako podstawa dla operatora systemu elektroenergetycznego do nagradzania poprawy gotowości elektrowni lub penalizacji jej spadku.

ZAHARIA SM. **The methodology of fatigue lifetime prediction and validation based on accelerated reliability testing of the rotor pitch links**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 638–644, http://dx.doi.org/10.17531/ein.2019.4.13.

Because the industrial products have lifetimes, without failing, of up to millions of cycles, it is mandatory that the aerospace field puts into practice the accelerated testing techniques. The lifetime prediction methodology for industrial products presented in this paper was put into practice by performing accelerated reliability testing on an aerospace product (the pitch link of a helicopter). The results showed a significant reduction of the testing time and costs. One important aspect highlighted in this paper is the equivalence between accelerated reliability testing and the traditional reliability testing, by using the two fundamental principles of the accelerated experiments: first, the stresses applied must not alter the physical mechanism through which the defects are produced and second, the conservation of the distribution laws of the failure times. In this way, by equivalence of the accelerated experiments, the methodology contained in this paper was validated.

ZAHARIA SM. **Metodologia prognozowania trwałości zmęczeniowej oraz jej walidacja w oparciu o przyspieszone badania niezawodności dźwigni skoku wirnika nośnego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 638–644, http://dx.doi.org/10.17531/ein.2019.4.13.

Ponieważ okresy bezawaryjnego użytkowania produktów przemysłowych stosowanych w w branży lotniczej mogą wynosić nawet kilka milionów cykli, badanie niezawodności tych wyrobów wymaga zastosowania technik badania przyspieszonego. Metodologię prognozowania czasu pracy produktów przemysłowych przedstawioną w niniejszym artykule wykorzystano w badaniach przyspieszonych niezawodności dźwigni skoku wirnika nośnego helikoptera. Wyniki wykazały, że proponowana metoda pozwala na znaczną redukcję czasu i kosztów badania. Ważnym aspektem, podkreślonym w niniejszej pracy, jest równoważność przyspieszonych i tradycyjnych badań niezawodności, którą można uzyskać respektując dwie podstawowe zasady eksperymentów przyspieszonych: po pierwsze, zastosowane naprężenia nie mogą zmieniać fizycznego mechanizmu, który prowadzi do powstania wady, a po drugie, należy przestrzegać praw dotyczących rozkładu czasów uszkodzeń. Przeprowadzone badania potwierdzają poprawność proponowanej metody.

KAMIŃSKI Z, RADZAJEWSKI P . **Calculations of the optimal distribution of brake force in agricultural vehicles categories R3 and R4**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 645–653, http://dx.doi.org/10.17531/ein.2019.4.14.

Fulfilling the requirements of the EU Directive 2015/68 in the area of braking for agricultural trailers depends on the proper selection of individual components of the braking system. This paper describes the requirements regarding braking performance and distribution of brake forces in agricultural trailers in R3 and R4 categories. On this basis, a methodology for calculating the optimal linear distribution of brake forces, characteristic for agricultural trailers with pneumatic braking systems, has been developed. The examples of calculation of an optimal distribution of brake forces for a two- and three-axle trailer with a tandem suspension system of the rear axle assembly have been provided. The optimization algorithm with the Monte Carlo method has been described, based on which a computer program was developed to select a linear distribution of brake forces in a three-axle trailer with 'walking beam' and 'bogie' suspensions. The presented calculations can be used in the design process to select the parameters of wheel braking mechanisms and then the characteristics of the pneumatic valves of the braking system.

KAMIŃSKI Z, RADZAJEWSKI P . **Obliczenia optymalnego rozdziału sił hamujących w przyczepach rolniczych kategorii R3 i R4**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 645–653, http://dx.doi.org/10.17531/ein.2019.4.14.

Spełnienie wymagań Dyrektywy UE 2015/68 w zakresie hamowania przyczep rolniczych zależy od właściwego doboru poszczególnych komponentów układu hamulcowego. W pracy opisano wymagania dotyczące skuteczności hamowania oraz rozdziału sił hamujących w przyczepach rolniczych kategorii R3 i R4. Na tej podstawie opracowano metodykę obliczeń optymalnego liniowego rozdziału sił hamujących, charakterystycznego dla przyczep rolniczych z pneumatycznymi układami hamulcowymi. Zamieszczono przykłady obliczeń optymalnego rozdziału sił hamujących dla przyczepy dwu i trzyosiowej z tandemowym układem zawieszenia zespołu osi tylnych. Opisano algorytm optymalizacji metodą Monte Carlo, na podstawie którego opracowano program komputerowy do doboru liniowego rozdziału sił hamujących w przyczepie trzyosiowej z zawieszeniem „walking beam" i „bogie". Przedstawione obliczenia można wykorzystać w procesie projektowania do doboru parametrów kołowych mechanizmów hamulcowych, a następnie charakterystyk zaworów pneumatycznych układu hamulcowego.

QUEZADA DEL VILLAR AV, RODRÍGUEZ-PICÓN LA, PÉREZ-OLGUÍN IJC, MÉNDEZ-GONZÁLEZ LC. **Stochastic modelling of the temperature increase in metal stampings with multiple stress variables and random effects for reliability assessment**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 654–661, http://dx.doi.org/10.17531/ein.2019.4.15.

Many products wear out over time even before they fail or stop working, therefore, through accelerated degradation tests one is able to make inferences about statistical

QUEZADA DEL VILLAR AV, RODRÍGUEZ-PICÓN LA, PÉREZ-OLGUÍN IJC, MÉNDEZ-GONZÁLEZ LC. **Ocena niezawodności z wykorzystaniem stochastycznego modelu wzrostu temperatury w metalowych wytłoczkach, uwzględniającego wielorakie zmienne naprężeniowe oraz efekty losowe**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 654–661, http://dx.doi.org/10.17531/ein.2019.4.15.

Wiele produktów zużywa się z upływem czasu zanim nawet ulegną uszkodzeniu lub przestaną działać. Badania przyspieszonego starzenia pozwalają wyciągać wnioski na

parameters or the distributions of a product useful life. Since many devices experience different types of variation due to unobservable factors during the manufacturing processes or under certain operating conditions; these situations lead to the need in developing accelerated degradation models with several variables of acceleration and random effects. The proposed model in this paper, is a model based on the gamma process with random effects to have a better analysis of degradation. This model is applied to the analysis of the temperature increase of metal stampings that are affected by multiple explanatory variables. In addition, a statistical inference method based on a Bayesian approach is used to estimate the unknown parameters to then perform a reliability analysis after obtaining the first-passage time distributions.

temat parametrów statystycznych lub rozkładów okresu użytkowania produktu. Wiele urządzeń podlega różnym rodzajom zmienności pod wpływem działania nieobserwowalnych czynników występujących podczas procesu produkcyjnego lub w pewnych warunkach pracy; sytuacje te wymagają opracowania modeli przyspieszonego starzenia uwzględniających wielorakie zmienne przyspieszenia oraz efekty losowe. Zaproponowany w przedstawionym artykule model opiera się na procesie gamma z efektami losowymi, dzięki czemu pozwala na lepszą analizę degradacji. Model ten zastosowano do analizy wzrostu temperatury w metalowych wytłoczkach, na które oddziałuje wiele zmiennych objaśniających. Ponadto do oszacowania nieznanych parametrów wykorzystano metodę wnioskowania statystycznego opartą na podejściu bayesowskim. Umożliwiło to analizę niezawodności po uzyskaniu rozkładów czasu pierwszego przejścia.

BORUCKA A, NIEWCZAS A, HASILOVA K. **Forecasting the readiness of special vehicles using the semi-Markov model**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 662–669, http://dx.doi.org/10.17531/ein.2019.4.16.

The vehicle exploitation system, consisting of statistically identical objects that perform intervention tasks, not subject to systematic changes, can be modelled as a stationary stochastic process. Such a model allows to determine the probabilistic indicators of current and boundary readiness of the system. This article presents the use of the semi-Markov process, based on three operating states: operation, ready-to-be-used and repair, to study a transport system consisting of special vehicles. On the example of a sample consisting of police patrol cars, experimental studies of the intensity of fleet utilization, time of failure-free operation of vehicles were carried out, and it was demonstrated that the examined transport system is characterized by a satisfactory, stationary readiness coefficient. The developmental possibilities of the presented modelling method were emphasized.

BORUCKA A, NIEWCZAS A, HASILOVA K. **Prognozowanie gotowości pojazdów specjalnych na podstawie modelu semi-Markowa**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 662–669, http://dx.doi.org/10.17531/ein.2019.4.16.

System eksploatacji samochodów, które realizują zadania interwencyjne, niepodlegający systematycznym zmianom może być modelowany jako stacjonarny proces stochastyczny. Taki model pozwala wyznaczyć probabilistyczne wskaźniki bieżącej i granicznej gotowości systemu. W niniejszym artykule, do modelowania systemu eksploatacji pojazdów specjalnych, wykorzystano proces semi-Markowa, oparty na trzech stanach eksploatacyjnych: użytkowania, postoju użytkowego i naprawy. Na przykładzie próby radiowozów policyjnych przeprowadzono doświadczalne badania intensywności użytkowania floty, czasu bezawaryjnej pracy pojazdów a także wykazano, że badany system transportowy charakteryzuje się zadowalającym, stacjonarnym współczynnikiem gotowości. Podkreślono rozwojowe możliwości przedstawionej metody modelowania.

ŻYWICA G, KACZMARCZYK TZ. **Experimental evaluation of the dynamic properties of an energy microturbine with defects in the rotating system**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 670–678, http://dx.doi.org/10.17531/ein.2019.4.17.

Today's energy systems increasingly use various types of microturbines to produce electricity. A specific feature of such machines is a high-speed rotor, whose rotational speed can be higher than 100,000 rpm. Failure-free operation of high-speed microturbine rotors requires both special design and high precision during the manufacturing process. What is more, proper procedures must be followed during run-up and coast-down phases; and also, dedicated diagnostic systems have to be used. This article discusses the experimental research conducted on a 2.5 kW vapour microturbine that operated in a prototypical combined heat and power plant. A series of measurements was carried out to evaluate the dynamic performance of the machine during normal operation. After the appearance of certain defects in the rotating system, it was necessary to perform a new series of measurements in order to assess the dynamic properties of the machine. The measurements results obtained in the form of vibration velocity spectrums made it possible to define diagnostic symptoms corresponding to particular defects. Similar diagnostic symptoms can occur during the operation of this class of turbomachines.

ŻYWICA G, KACZMARCZYK TZ. **Eksperymentalna ocena właściwości dynamicznych mikroturbiny energetycznej w obecności defektów układu wirującego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 670–678, http://dx.doi.org/10.17531/ein.2019.4.17.

We współczesnych systemach energetycznych coraz częściej do wytwarzania energii elektrycznej stosowane są różnego typu mikroturbiny. Charakterystyczną cechą takich maszyn są wysokoobrotowe wirniki, których prędkości obrotowe mogą przekraczać nawet 100 000 obr/min. Praca wirnika w takich warunkach wymaga zastosowania specjalnych rozwiązań konstrukcyjnych i bardzo dużej precyzji wykonania, a podczas eksploatacji zachowania odpowiednich procedur przy rozruchu i odstawieniu, a także stosowania dedykowanych systemów diagnostycznych. W niniejszym artykule zostały omówione badania eksperymentalne mikroturbiny parowej o mocy 2,5 kW, pracującej w prototypowym układzie kogeneracyjnym. Wykonane pomiary obejmowały ocenę stanu dynamicznego podczas normalnej pracy maszyny oraz badania jej właściwości dynamicznych w obecności defektów układu wirującego. Uzyskane wyniki pomiarów, w postaci rozkładów częstotliwościowych drgań, pozwalają na zdefiniowanie symptomów diagnostycznych typowych dla różnych defektów, które mogą pojawić się podczas eksploatacji tej klasy maszyn wirnikowych.

KOZŁOWSKI E, MAZURKIEWICZ D, ŻABIŃSKI T, PRUCNAL S, SĘP J. **Assessment model of cutting tool condition for real-time supervision system**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 679–685, http://dx.doi.org/10.17531/ein.2019.4.18.

Further development of manufacturing technology, in particular machining requires the search for new innovative technological solutions. This applies in particular to the advanced processing of measurement data from diagnostic and monitoring systems. The increasing amount of data collected by the embedded measurement systems requires development of effective analytical tools to efficiently transform the data into knowledge and implement autonomous machine tools of the future. This issue is of particular importance to assess the condition of the tool and predict its durability, which are crucial for reliability and quality of the manufacturing process. Therefore, a mathematical model was developed to enable effective, real-time classification of the cutting blade status. The model was verified based on real measurement data from an industrial machine tool.

KOZŁOWSKI E, MAZURKIEWICZ D, ŻABIŃSKI T, PRUCNAL S, SĘP J. **Model oceny stanu narzędzia skrawającego dla systemu nadzoru w czasie rzeczywistym**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 679–685, http://dx.doi.org/10.17531/ein.2019.4.18.

Dalszy rozwój inżynierii produkcji, w szczególności obróbki skrawaniem, wymaga poszukiwania nowych innowacyjnych rozwiązań technologicznych. Dotyczy to w szczególności zaawansowanego przetwarzania danych pomiarowych pochodzących z systemów diagnostycznych i monitorujących. Rosnąca ilość danych gromadzonych przez wbudowane systemy pomiarowe wymaga opracowania skutecznych narzędzi analitycznych, aby efektywnie przekształcać dane w wiedzę i wdrażać autonomiczne obrabiarki przyszłości. Kwestia ta ma szczególne znaczenie dla oceny stanu narzędzia i przewidywania jego trwałości, które są kluczowe dla niezawodności i jakości procesu produkcyjnego. Dlatego opracowano nowy model matematyczny, którego zadaniem jest skuteczna klasyfikacja stanu ostrza narzędzia skrawającego realizowana w czasie rzeczywistym. Opracowany model został zweryfikowany na podstawie rzeczywistych danych pomiarowych z przemysłowej obrabiarki.

GEVORKYAN E, PROKOPIV M, RUCKI M, MOROZOW D. **Durability and exploitation performance of cutting tools made out of chromium oxide nanocomposite materials**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 686–691, http://dx.doi.org/10.17531/ein.2019.4.19.

This article is devoted to nanoscale composite materials based on Cr_2O_3 obtained by the activated electric fields sintering procedure. In the paper, exploitative properties of the sintered system of Cr_2O_3 – AlN nanocomposite was examined. Mechanical properties of the material were examined, especially from the perspective of its performance in the cutting tools. In particular, its wear was tested at different cutting speeds, as well as for intermittent hard cutting, and the results were compared with other materials available in the market. Compared to other cutting tools of the

GEVORKYAN E, PROKOPIV M, RUCKI M, MOROZOW D. **Trwałość i właściwości eksploatacyjne narzędzi skrawających wykonanych z nanokompozytu tlenku chromu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 686–691, http://dx.doi.org/10.17531/ein.2019.4.19.

Artykuł jest poświęcony właściwościom eksploatacyjnym materiałów kompozytowych na bazie Cr_2O_3 wytworzonych metodą spiekania w polu elektrycznym. W szczególności poświęcono uwagę nanokompozytowemu spiekowi Cr_2O_3 – AlN wykorzystywanemu do wytwarzania narzędzi skrawających. Zbadano właściwości mechaniczne materiału z uwzględnieniem trwałości ostrzy i powierzchni skrawających. Zbadano zużycie przy różnych prędkościach skrawania w warunkach ciągłych i przerywanych. W porównaniu do ostrzy podobnej klasy, np. Bichromit-R, badane płytki wykazywały podobną trwałość

same class, Bichromit-R performed the same lifetime for 3-5 times higher cutting speeds, or up to 45% longer lifetime for the same cutting speed. The results lead to the conclusion that composite nanostructure improves substantially exploitation characteristics of the cutting tools.

przy wyższych 3 do 5 razy prędkościach skrawania, albo pracowały ok. 45% dłużej przy tych samych prędkościach. Wyniki badań prowadzą do wniosku, że nanostruktura materiału kompozytowego znacząco polepsza właściwości eksploatacyjne ostrzy skrawających.

WALISZYN A, ADAMKIEWICZ A. **Studies on resistance to erosion of nickel and its alloys to be used in elements of fluid - flow machines**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 692–696, http://dx.doi.org/10.17531/ein.2019.4.20.

The article presents results of studies on metal resistance to erosive damage taking place under the influence of hydraulic cavitation. On the basis of earlier research, a hypothesis on fatigue character of erosive wear and a dependence of metal resistance to erosive damage on its crystalline lattice structure has been assumed. To verify this hypothesis, metals with different crystalline lattice structures like steel 45 (flat-centred structure), nickel 200/201 and nickel alloy Monel 400 (hexagonal structure) have been tested at a cavitation-strike stand. Results obtained there confirmed the assumed hypothesis, at the same time justifying the use of nickel protective coatings in fluid-flow machines.

WALISZYN A, ADAMKIEWICZ A. **Badania odporności erozyjnej niklu i jego stopu do zastosowania w elementach maszyn przepływowych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 692–696, http://dx.doi.org/10.17531/ein.2019.4.20.

W artykule przedstawiono wyniki badań odporności metali na uszkodzenia erozyjne zachodzące pod wpływem kawitacji hydraulicznej. Na podstawie wyników wcześniejszych badań, przyjęto hipotezę o zmęczeniowym charakterze zużycia erozyjnego oraz zależności odporności metali na zniszczenia erozyjne od struktury ich sieci krystalicznej. Dla potwierdzenia przyjętej hipotezy na stanowisku kawitacyjno-udarowym sprawdzono metale z różnymi sieciami krystalicznymi: stal 45 (sieć płasko centralna), nikiel 200/201 oraz stop niklu Monel 400 (sieć heksagonalna). Otrzymane wyniki badań potwierdziły przyjętą hipotezę, wskazując tym samym na zasadność stosowania niklowych powłok ochronnych w maszynach przepływowych.

Article citation info:
HAN T, JIANG D. Fault diagnosis of multistage centrifugal pump unit using non-local means-based vibration signal denoising. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 539–545, http://dx.doi.org/10.17531/ein.2019.4.1.

Te HAN
Dongxiang JIANG

FAULT DIAGNOSIS OF MULTISTAGE CENTRIFUGAL PUMP UNIT USING NON-LOCAL MEANS-BASED VIBRATION SIGNAL DENOISING

DIAGNOZOWANIE USZKODZEŃ WIELOSTOPNIOWEJ POMPY ODŚRODKOWEJ Z WYKORZYSTANIEM METODY ODSZUMIANIA SYGNAŁU DRGAŃ W OPARCIU O ŚREDNIE NIELOKALNE

In real industry environment, the signal characteristics of multistage centrifugal pump vibration signal are easily submerged by strong background noise. To settle this problem, the nonlocal means (NLM) approach is proposed for the denoising of multistage centrifugal pump in this paper. Utilizing the similarity theory, the NLM method has achieved a wide range of applications in the fields of image processing and biomedical signal denoising. Due to the periodic characteristics and redundancy, NLM is successfully applied to the de-noising of 1-D machinery vibration signal. The numerical simulation experiments with different SNRs verify the effectiveness and the superiority of the proposed method. Besides, the selection principles of core parameters in NLM are discussed. The real engineering cases analysis demonstrates that the NLM can effectively filter out the background noise and realize the weak fault feature enhancements. The proposed noise reduction method is superior to traditional wavelet coefficient method.

***Keywords**: vibration signal denoising, multistage centrifugal pump, nonlocal means, diagnosis.*

W rzeczywistym środowisku przemysłowym, charakterystyki sygnału drgań wielostopniowej pompy odśrodkowej są zagłuszane przez silny szum tła. Problem ten można rozwiązać stosując zaproponowane w niniejszej pracy podejście oparte na algorytmie średnich nielokalnych (non-local means, NLM). Wykorzystująca teorię podobieństwa metoda NLM znajduje szeroki zakres zastosowań w dziedzinie przetwarzania obrazu i odszumiania sygnałów biomedycznych. Dzięki okresowemu charakterowi i redundancji sygnałów, NLM można z powodzeniem stosować do usuwania szumu jednowymiarowego sygnału drgań maszyn. Skuteczność proponowanej metody i jej przewagę nad stosowanymi dotychczas rozwiązaniami zweryfikowano na podstawie eksperymentów symulacyjnych z uwzględnieniem różnych stosunków sygnału do szumu (SNR). Ponadto omówiono zasady wyboru podstawowych parametrów NLM. Analiza przypadków inżynierskich pokazuje, że NLM pozwala skutecznie odfiltrowywać szumy tła i wzmacniać słabe symptomy akustyczne uszkodzenia. Proponowana metoda redukcji szumów przewyższa tradycyjną metodę współczynnika falkowego.

***Słowa kluczowe**: odszumianie sygnału drgań, wielostopniowa pompa odśrodkowa, średnie nielokalne, diagnozowanie.*

1. Introduction

Multistage centrifugal pump units are widely used in diverse areas, such as condensate water supply for nuclear power plants, drainage of submarine, agricultural irrigation, etc [7]. The condition monitoring and fault diagnosis technologies are significantly important to improve the reliability and stability during the operation of pump units [4, 9, 12, 13, 26]. Nowadays, vibration signal-based analysis is one of the most efficient methods for condition monitoring and fault diagnosis [6, 17, 19]. However, due to the interference of background noise, the obtained signals often cannot accurately reflect the running state of the pump units. The early fault characteristics are easily submerged by noise, lead to the inexact assessment of health conditions. Consequently, fast and effective noise reduction method is of great significance to accurately extract fault features in the condition monitoring and fault diagnosis of pump units [1, 8].

Because of the complex structure of the multistage centrifugal pump, multiple internal excitation sources, the vibration signal is characterized by strong non-linearity and non-stationarity [1]. Traditional Fourier transform-based denoising methods and FIR filtering methods can't process the non-stationary signal well. Most of current researches about vibration signal denoising focus on some popular techniques, mainly including wavelet analysis, empirical mode decomposition (EMD), singular value decomposition (SVD), etc. Wavelet transform methods have the capability of analyzing non-stationary signals [5, 29]. By scaling and shifting the wavelet basis functions, the local features of signal in time-frequency domain can be appropriately extracted. The wavelet thresholding-based denoising methods have been widely used in noised reduction and diagnosis of mechanical signals. The popular threshold selection methods mainly include uniform threshold, unbiased estimation threshold, minimax threshold, heuris-

tic threshold, etc. Yang et al. proposed an improved wavelet adjacent coefficient-based denoising method by setting the adjacent wavelet coefficient as the whole threshold value, so that filtering the noise and efficiently retaining the impact characteristics in the raw signal [24]. Wu et al. used the threshold estimation method for overlapping blocks to improve the effect of noise reduction [23]. Because of the complexity of mechanical signals, wavelet analysis methods need to set wavelet basis according to signal characteristics in advance. At the same time, the number of decomposition layers has a great influence on the effect of noise reduction. EMD based denoising methods are another branch with no need to set predefined basis function. By decomposing the original signal into a series of Intrinsic Mode functions (IMFs) from high frequency to low frequency, the IMFs that contain the main information can be selected and reconstructed, the high frequency noise in other IMFs can be filtering. Zhang et al. utilized the EMD to decompose the current signal and extract blade imbalance fault feature, which is concealed by the supply frequency and the environment noise [14]. However, end effects, modal aliasing and the selection criteria of IMFs are still the challenges in this branch [1, 10]. The scholars have proposed a series of improved algorithms to overcome these drawbacks, such as ensemble empirical mode decomposition (EEMD) [10], local mean decomposition (LMD) [29] and variational mode decomposition (VMD) [11], while increased the complexity of denoising procedure. In addition, by discarding the high frequency components that represent noise, useful high frequency information is also discarded. The signal processing for each IMFs in such approached is therefore another difficulty. SVD based denoising mainly reconstruct the original signal into a Hankel matrix, and then the matrix is transformed into a factorization form, where diagonal entries are known as the singular values. Each singular value can be reconstructed into a signal components. Generally, the large values of front singular value represent the useful information of raw signal, while the small values of the last ones are related to random noise [20]. The works about the methods to select effective singular values have been often reported [11, 18, 20].

Overall, these aforementioned methods are all based on the idea of decomposition, which may suffer from two difficulties in practical engineering. (1) The selection of predefined basis function and parameter design largely rely on priori knowledge and expert experience. (2) After decomposition, much efforts are need to select and process the useful components while discard the components containing noise. It is highly desirable to investigate the vibration signal denoising methods with easy implementation and simple parameter determination. Non-local mean (NLM) algorithm is a denoising method proposed by Buades et al. in the field of image processing by taking advantage of similar structure characteristics of image blocks [2]. Different from the idea of signal decomposition, the non-local mean algorithm searches for similar structures in the region, and then removes noise by weighted average. Due to its excellent noise reduction ability and relatively simple parameter selection, this algorithm has been successfully applied in image processing, denoising of biomedical signals and other fields [15, 21, 25, 28]. For rotating machinery, vibration signals are usually periodic and cyclo-stationary, and the signal mode reflecting the characteristics of equipment state appears repeatedly with abundant redundant information. These characteristics provide the foundation for the application of this algorithm in denoising of 1-dimensional mechanical vibration signal [27]. This work applied the NLM methods to the denoising of vibration signal of multistage centrifugal pump. The validity and superiority of the method are verified by simulation experiment and analysis of vibration data in real pump unit, providing a fast and effective denoising method for practical engineering.

2. The principle of NLM filtering

NLM filtering focuses on denoising in each region. The core idea is to search similar signal blocks in a neighborhood class and conduct weighted average so as to remove noise and large amount of redundant information in the raw signal. For the vibration signals of rotating machinery, some signal characteristic modes, such as pulse characteristics and periodically appeared rotating frequency components, generally exist. While the noise superimposed on the signal block is random distribution, which can be effectively filtered by weighted average. The core problem of non-local mean denoising is to recover the original signal from a signal containing additive noise. The signal model with additive noise is as follows:

$$v = u + n \tag{1}$$

where v is the measured signal, u is the theoretically noiseless signal and n is the additive white gaussian noise. Given a measured sample s, $\hat{u}(s)$ represents the estimation of the original noise-free signal and it can be calculated by searching a series of similar signal blocks in a neighborhood and making weighted averages. The formula can be described as follows:

$$\hat{u}(s) = \frac{1}{Z(s)} \sum_{t \in N(s)} w(s,t) v(t) \tag{2}$$

where $w(s,t)$ represents the weight of similarity between the t-center signal block and the s-center target signal block, $N(s)$ means the searching neighborhood centered on the target signal block, $Z(s) = \sum_t w(s,t)$ is a normalization constant to represent the sum of the weights of all similar blocks. The weight $w(s,t)$ is calculated as:

$$w(s,t) = \exp\left(-\frac{\sum_{\delta \in \Delta} \left(v(s+\delta) - v(t+\delta) \right)^2}{2 L_\Delta \lambda^2} \right) \tag{3}$$

where λ the parameter of filter, L_Δ represents the data points for signal blocks. The similarity is estimated by the Euclidean distance between the t-center signal block and the s-center target signal block. In Eq (3), the weight of each signal block itself $w(s,s) = 1$. In the field of image processing, to obtain better smoothing effect, the formula for calculating the similarity weight of the central signal block is as follows:

$$w(s,s) = \max_{t \in N(s),\ t \neq s} w(s,t) \tag{4}$$

Considering the non-stationary characteristics of mechanical signals, there is still a large difference in amplitude between the signal blocks with similar structures, which may lead to a low similarity weight and further smooth the original signal. Excessive smooth can result in the loss of signal detail. Therefore, the similar weight correction of the central signal block is not adopted in this work.

It can be seen that the weight depends on the similarity between the signal blocks, rather than only considering the center distance between the signal blocks, so that the weighted average denoising can retain the signal details to the maximum extent. When the search neighborhood covers the whole signal, the algorithm realizes the true non-local mean. However, computational burden increases linearly with the signal length. For one dimensional signal with N data points and search radius with M data points, the computational complexity

is $O(L_{\Delta}NM)$. In this work, the fast non-local mean method proposed by Darbon et al. is adopted for subsequent research [3]. This method can accelerate the calculation process of similar weights by reducing nested cycles, and the computational complexity after optimization is $O(2NM)$. Fig. 1 gives the illustration of NLM parameters. The red patch with center s means the target signal patch, while the yellow patch with center t represents the searched patch in a neighbourhood. The searching region for the neighbourhood contains 2M points with the center of s.

Essentially, different from signal decomposition in most of methods, such as wavelet and EMD, NLM is based on the idea of statistical neighbourhood filter. In contrast to the methods, such as wavelets and Fourier transform, which require predefined basis function, NLM reduces the noise in the raw signal by approximating the signal patch with the self-similar patches in the original signal, instead of decomposition with basis function. Moreover, NLM avoids to identify which decomposed components represent dominantly noise and which components contain primarily main information in the methods, such as EMD and SVD. Consequently, the basic principle of NLM filtering shows the application prospect for vibration signal denoising.

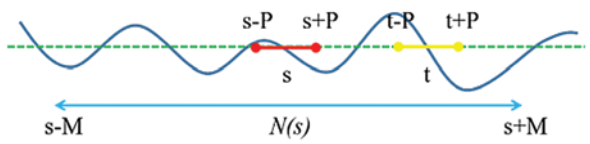


Fig. 1. Illustration of NLM parameters.

3. Parameters criteria

The length of signal block (2P+1) is an important parameter in NLM denoising. When the value of P is too small, the signal block cannot reflect the typical characteristic mode of the signal. At the same time, it is seriously disturbed by noise. When the P is set to a large value, the signal block contains too much information, which is easy to cause smoothing effect in the weighted evaluation process. For the one-dimensional mechanical vibration signal, an appropriate P is usually half of the typical characteristic length, and the specific typical signal block length needs to be selected according to the characteristics of the signal.

For the search radius M of the neighborhood, theoretically, a larger M is benefit to find more similar signal blocks, and more redundant information will be enriched to achieve better noise reduction effect, however, it will bring significant increase in the amount of calculation.

The filter bandwidth λ primary control the smoothing effect in denoising process. A too small λ is likely to cause noise disturbance and affect the size of the weight between the signal blocks, causing an inadequate average. A too large λ will cause a large similarity weight between low-similarity signal blocks, resulting in excessive smoothness and loss of local detail characteristics of the signal. In the field of image processing, according to the SURE criteria of Ville and Kocher [22], $\lambda = 0.5\sigma$, where σ is the standard deviation of noise. Referring to the application of NLM in one-dimensional signal denoising [21], λ is set to 0.6σ in this work. For the real signal, the noise variance can be estimated based on wavelet coefficient.

4. Numerical analysis

Due to mechanical and water flow excitation, vibration signal of multistage centrifugal pump is mainly composed of rotating frequency and its harmonic frequency components. For bearing faults, the signal has obvious impact characteristics. To verify the effectiveness of the method, two sets of simulation signals are designed and analyzed. The noise reduction effect is quantitatively evaluated by introducing three indexes of SNR improvement, mean square error (MSE) and distortion rate (DR). The calculation formulas are as follows:

$$SNR_{imp} = 10\log_{10}\frac{\sum_{n=1}^{N}(v[n]-u[n])^2}{\sum_{n=1}^{N}(\hat{u}[n]-u[n])^2} \tag{5}$$

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(\hat{u}[n]-u[n])^2 \tag{6}$$

$$DR = 100\sqrt{\frac{\sum_{n=1}^{N}(\hat{u}[n]-u[n])^2}{\sum_{n=1}^{N}u^2[n]}} \tag{7}$$

where N means the length of signal and the other symbols are same as stated above. The first set of simulation signals is designed as follows:

$$x(t) = \sum_{i=1}^{6} A_i sin 2\pi f_i t \tag{8}$$

where $A_1 - A_6$ are set to 20, 4.5, 2.55, 1.5, 0.4, and 0.3 respectively, $f_1 - f_6$ are set to 20, 2×20, 3×20, 4×20, 0.2×20, 0.3×20 respectively, the sampling frequency is 1000Hz and signal length is 2048 points. The white gaussian noise is added and the SNR is 8dB. The time-domain waveforms of the simulated signal and the noisy signal are shown in Fig. 2:

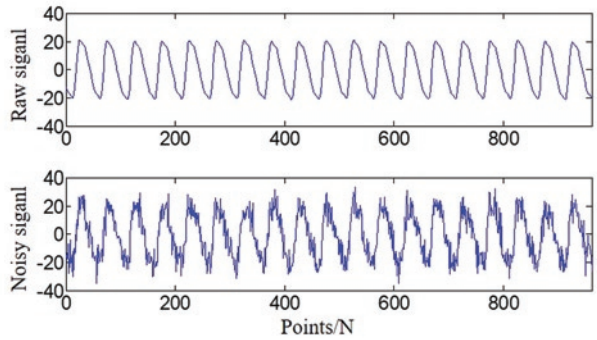


Fig. 2. Time waveform of simulation signal 1

NLM and wavelet soft threshold method (Wden) are used for noise reduction and comparison. The detailed parameters of the two methods are set as follows. NLM: p=30, M=1000, λ=0.6σ. Wden: db10 is selected as wavelet base function, decomposition layer is 5, heuristic threshold criterion is used. The denoising signals are shown in Fig. 3. It can be seen that the NLM algorithm performs better in filtering noise and reflecting the characteristics of the original waveform. The wavelet threshold denoising can obtain smoother original waveform. However, the wavelet basis structure needs to meet the orthogonal premise, while the wavelet basis function insufficiently matches the original waveform features, resulting in the phenomenon of local distortion in denoising signal. Furthermore, the above three indicators are used to conduct quantitative evaluation on the noise reduction performance under different SNR conditions, as listed in Table 1.

The second group of simulation signals are periodic shock signals, and the calculation formula is as follows:

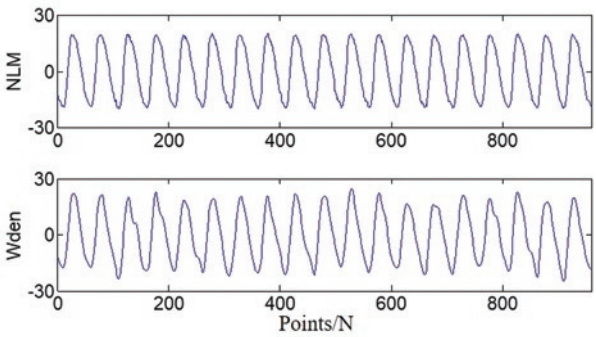


Fig. 3. The comparison of denoising results of two methods for simulation signal 1

Table 1. The denoising results of two methods with different SNRs for simulation signal 1

SNR/dB	NLM			Wden		
	SNR_{imp}	MSE	DR/ 10^3	SNR_{imp}	MSE	DR/ 10^3
8	13.33	1.58	0.47	7.34	6.26	1.43
7	13.73	1.75	0.59	7.48	7.39	1.71
6	11.76	3.58	0.60	8.33	7.88	1.79
5	12.62	3.70	0.75	7.42	12.25	2.26
4	12.43	4.85	0.98	8.36	12.40	2.46
3	10.15	10.22	1.90	7.21	20.09	3.24

$$x(t) = \exp^{-at} \sin(2\pi \times f_c nT)$$
$$t = mod\left(nT, \frac{1}{f_m}\right) \tag{9}$$

where $a = 150$, carrier frequency $f_c = 1000$, fault characteristic frequency $f_m = 20$, sampling frequency is 2500Hz and signal length is 2048 points. The SNR of the noise signal is 4dB, and the time-domain waveform is shown in Fig. 4. Using the same parameter settings above, the results of the two denoising methods are shown in Fig. 5. Under different SNR conditions, the quantitative evaluation results of noise reduction are given in Table 2. Similar conclusions can be drawn that NLM denoising not only enhances the shock characteristics of the signal, but also better guarantees the waveform characteristics of the original impulses. However, the classical wavelet threshold method has obvious distortion. NLM noise reduction signal has a higher reduction degree, compared to the original signal.

Combined with the second set of simulation signal, two key parameters P and M in the NLM denoising process reduction are discussed to provide a clearer guiding principle for the parameter selection in the practical tasks. The curves of SNR enhancement and MSE along with parameters are shown in Fig. 6. Since the result of DR is similar to that of the MSE, it is not shown here. It can be seen that a best noise reduction effect is achieved, when parameter P is set between 30 and 50 points. In the simulation signals, each impulse feature lasts for about 100 points. When P is between 30-50, the signal block with length (2P+1) just reflects the shock feature of the original signal, and thus achieves a good noise reduction effect. At the same time, with the increase of M length in the search neighborhood, the denoising effect is also improved correspondingly. When M contains enough points (M > 2048), the lifting speed of each index gradually slows down.

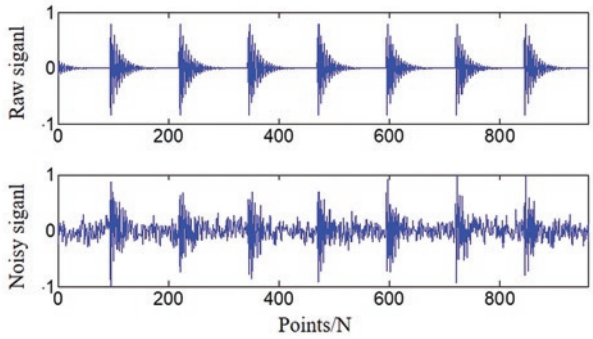


Fig. 4. Time waveform of simulation signal 2

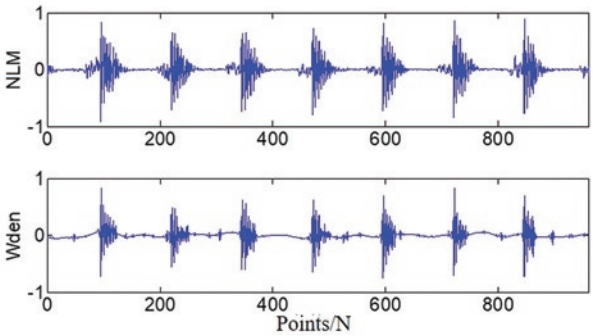


Fig. 5. The comparison of denoising results of simulation signal 2 for two methods

Table 2. The denoising results of two methods with different SNRs for simulation signal 2

SNR/dB	NLM			Wden		
	SNR_{imp}	MSE/ 10^{-3}	DR	SNR_{imp}	MSE/ 10^{-3}	DR
8	10.3	0.51	13.56	2.46	3.1	65.75
7	9.97	0.68	17.25	2.91	3.5	80.62
6	9.73	0.91	23.13	3.40	3.9	99.32
5	8.50	1.5	30.73	3.35	5.0	105.6
4	8.19	2.1	60.96	3.68	5.9	151.7
3	5.93	4.4	103.8	4.02	6.8	157.7

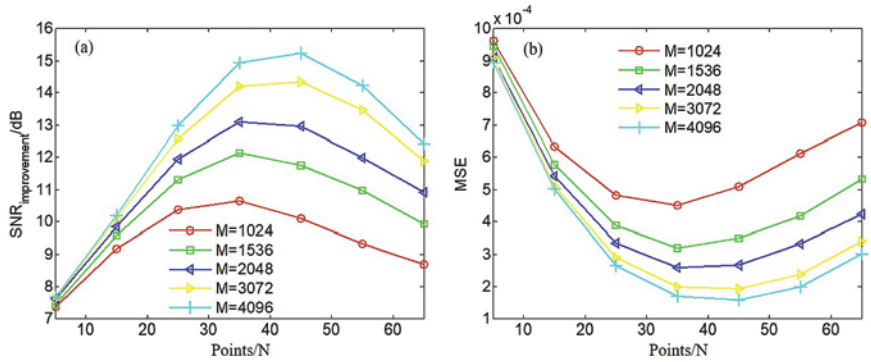


Fig. 6. The improvement of evaluation indicators versus P and M($\lambda = 0.6\sigma$): (a) SNR improvement, (b) MSE

5. Case studies on vibration signal of multistage centrifugal pumps

In order to verify the effectiveness of NLM method in actual industrial tasks, this paper selects two cases of practical vibration fault diagnosis of multistage centrifugal pump sets for analysis. The first analysis case is that the misalignment fault of the motor and the centrifugal pump coupling leads to the excessively high vibration index. The rotating speed of the centrifugal pump is 1490r/min. Vibration signals are collected through the acceleration sensor located at the pump driving end with a sampling frequency of 1280Hz. Time-domain waveforms are shown in Fig. 7, which are further used for denoising (containing 1792 data points). It can be intuitively seen that the vibration signal contains very obvious rotating frequency component. At the same time, some high-frequency components are superimposed on the rotating frequency signal, which interferes with the further identification of faults. NLM is used for denoising. The parameters settings are p=10, M=1024, λ=0.6σ, the noise variance can be estimated from the high-frequency coefficient of the first layer in wavelet decomposition. wavelet soft threshold denoising method is also used for comparative analysis. Wavelet basis function is db10, which highly resembles the vibration signal no matter for the main frequency component or the impulses with high frequencies. In addition, db10 wavelet has orthogonal property which enables perfect reconstruction of signal and have been reported to produce best result according to related works [16]. The number of decomposition layer is 5, heuristic threshold criterion is used.

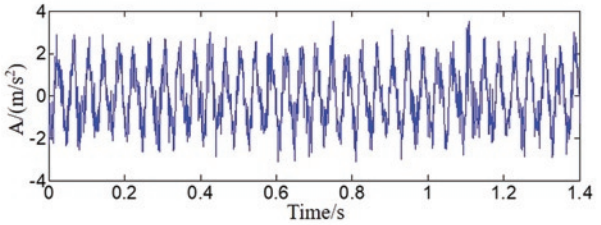


Fig. 7. The vibration signal of multistage centrifugal pump with misalignment

The denoising signals are shown in Fig. 8. From the NLM denoising results, it can be seen that the high frequency noise is effectively filtered. There are obvious harmonic frequency components in the signal. But strong noise interference still exists in the denoising signal for the wavelet threshold method. For practical engineering tasks, the selection of wavelet threshold has a great influence on the result of noise reduction.

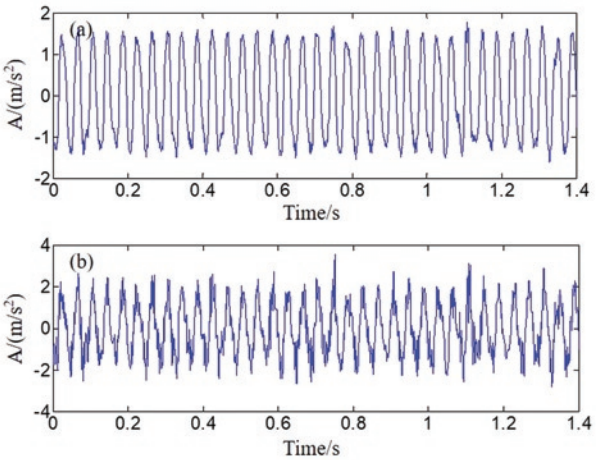


Fig. 8. The denoising results of for two methods in case study 1: (a) NLM, (b) Wden

The fault characteristics can be preliminarily detected from the harmonic frequency. To remove the main frequency component, namely the rotating frequency, and present a clear fault symptom, the ensemble empirical mode decomposition (EEMD) is further utilized to decompose denoising signal and extract the components with fault information for spectrum analysis. The results are shown in Fig. 9 The 1X, 2X and 3X frequency components can be accurately extracted from the spectrum of NLM method. The preliminary judgment is that the coupling alignment occurs. In the wavelet threshold noise reduction method, due to the interference of residual noise, the harmonic frequency components in the spectrum are not obvious.

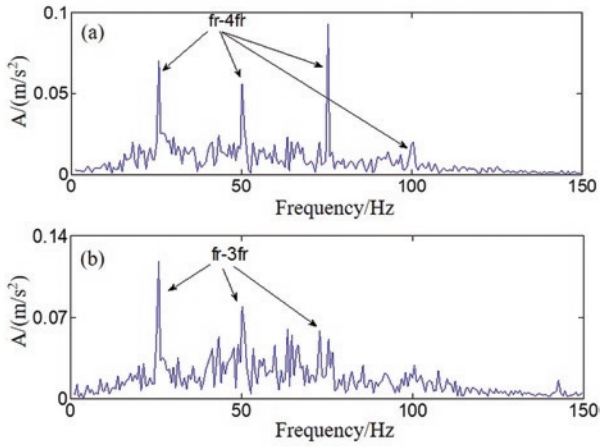


Fig. 9. The frequency spectra of fault component: (a) NLM and (b) Wden

The second case is a bearing fault vibration signal analysis in centrifugal pump. The faults occurs in the outer ring of non-driving end conical roller bearing, resulting in high vibration index. The rotating speed of the centrifugal pump is 1490r/min, and the fault characteristic frequency of the bearing outer ring is $f_o = 236.67$ Hz. The sampling frequency is 1280 Hz. The vibration waveform collected by the acceleration sensor at the non-driving end is shown in Fig. 10.

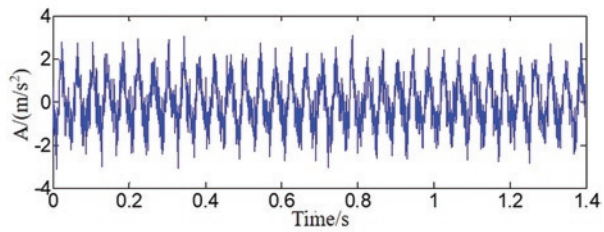


Fig. 10. The vibration signal of multistage centrifugal pump with bearing fault

Similar to case 1, the rotating frequency component in the signal can be clearly found, and at the same time, there are a large number of high-frequency components in the signal. NLM and wavelet threshold are used for denoising respectively. The parameter settings are the same as above. The results of NLM are shown in Fig. 11. A large number of high-frequency components are effectively filtered out, but some high-frequency impulse components still exist in the waveform after denoising. The local amplification of the denoising signal shows that these high-frequency components show obvious characteristics of periodic shock. The preliminary judgment is related to bearing failure. After the envelope analysis of the denoising signal, as shown in the Fig. 11(c), the double fault characteristic frequency $f_o = 473.3$ Hz can be exactly detected. Owing to the weighted average between similar signal blocks, it can be seen that NLM method can effectively filter out noise and enhance the periodicity similar signal

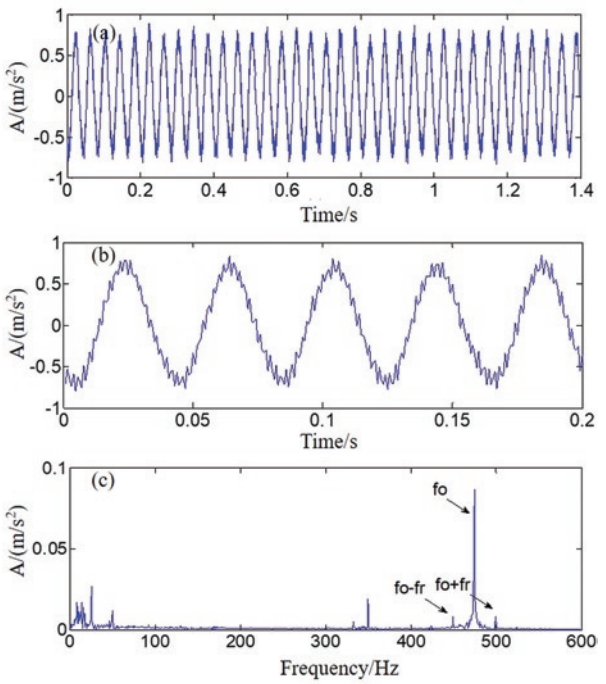


Fig. 11. The analysis of denoising results of NLM in case study 2: (a) denoising signal, (b) local amplification of the denoising signal and (c) envelope spectrum of denoising signal

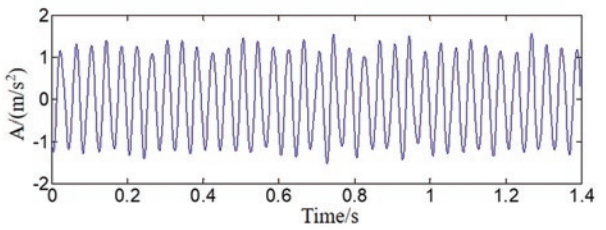


Fig. 12. The results of Wden in case study 2

components, providing convenience for rapid and accurate fault diagnosis. For comparison, the denoising waveform of wavelet threshold methods is shown in Fig. 12. Due to the problem of threshold selection, all the high frequency fault information is filtered out, only leaving smooth rotating frequency component, which cannot provide effective diagnostic information.

The effectiveness of NLM can be verified by the two case studies in real multistage centrifugal pump unit. Although this work only focuses on multistage centrifugal pump unit, the feature modes and characteristics are similar to other rotating machineries, such as wind turbine and stream turbine. For rotor-related faults, the harmonic frequencies components are is superimposed upon the rotating frequency component. For bearing and gear faults, the impulses and resonance components in high frequency are generally induced. Overall, the fault features appear periodically, and a large amount of similar information exist in the vibration signal of rotating machineries. Consequently, by comparing the local neighborhoods, finding the similar features on the basis of self-similarities, and removing the redundancy of similar patches, NLM is significantly suitable for analyzing the vibration signals of rotating machinery. Moreover, different from the decomposition-based denoising methods, there are no need of selecting basis function and setting complex decomposed parameters in NLM, providing an accessible denoising approach for practical application.

6. Conclusions

Noise reduction of mechanical vibration signal is the key and difficult point to realize accurate fault diagnosis. Considering the non-stationary characteristics of vibration signal and the difficulty of selecting the threshold value in wavelet denoising methods, this paper investigates vibration signal denoising of multistage centrifugal pump unit using NLM. Different from the traditional denoising idea of basis function decomposition, NLM uses the weighted average of similar signal blocks in a neighborhood, which has strong non-stationary signal processing ability and excellent adaptability. Moreover, this method has a wide application prospect because of its simple principle and easy parameter selection. The validity and superiority of the method are verified by the simulation signal analysis. The key parameter selections are also discussed. The analysis results of practical diagnosis cases of multistage centrifugal pump show that NLM can effectively filter the noise components with poor correlation between signal blocks and enhance the periodic fault characteristics.

Acknowledgement

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by the National Natural Science Foundation of China (No. 51174273) and the project of State Key Lab of Power Systems (No. SKLD16Z12).

References

1. An X and Yang J. Denoising of hydropower unit vibration signal based on variational mode decomposition and approximate entropy. Transactions of the Institute of Measurement and Control 2016; 38: 282-92, https://doi.org/10.1177/0142331215592064.
2. Buades A, Coll B, Morel J M. A Non-Local Algorithm for Image Denoising[C]// Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on IEEE 2005; 2 :60-65.
3. Darbon J, Cunha A, Chan T F. Fast nonlocal filtering applied to electron cryomicroscopy in IEEE International Symposium on Biomedical Imaging: From Nano To Macro. IEEE 2008:1331-1334, https://doi.org/10.1109/ISBI.2008.4541250.
4. Duan R, Lin Y, Zeng Y. Fault diagnosis for complex systems based on reliability analysis and sensors data considering epistemic uncertainty. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 558-566, https://doi.org/10.17531/ein.2018.4.7.
5. Guo-dong Yue, Xiu-shi Cui, Yuan-yuan Zou, Xiao-tian Bai, Yu-Hou Wu, Huai-tao Shi. A Bayesian wavelet packet denoising criterion for mechanical signal with non-Gaussian characteristic. Measurement 2019, 138: 702-712, https://doi.org/10.1016/j.measurement.2019.02.066.
6. Han T, Jiang DX, Zhao Q, Wang L and Yin K. Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. Transactions of the Institute of Measurement and Control 2018; 40: 2681-2693, https://doi.org/10.1177/0142331217708242.
7. Hernandezsolis A, Carlsson F. Diagnosis of Submersible Centrifugal Pumps: A Motor Current and Power Signature Approaches. Epe Journal European Power Electronics & Drives 2010; 20(1): 58-64, https://doi.org/10.1080/09398368.2010.11463749.

8. Junchao Guo, Dong Zhen, Haiyang Li, Zhanqun Shi, Fengshou Gu, Andrew.D. Ball. Fault feature extraction for rolling element bearing diagnosis based on a multi-stage noise reduction method. Measurement 2019; 139: 226-235, https://doi.org/10.1016/j.measurement.2019.02.072.
9. Kaluer S, Fekete K, Jozsa L, Klai Z. Fault diagnosis and identification in the distribution network using the fuzzy expert system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 621-629, https://doi.org/10.17531/ein.2018.4.13.
10. Kai Zheng, Jiufei Luo, Yi Zhang, Tianliang Li, Jiafu Wen, Hong Xiao. Incipient fault detection of rolling bearing using maximum autocorrelation impulse harmonic to noise deconvolution and parameter optimized fast EEMD. ISA Transactions 2019; 89: 256-271, https://doi.org/10.1016/j.isatra.2018.12.020.
11. Li, H., Bao, T., Gu, C., Chen, B. Vibration feature extraction based on the improved variational mode decomposition and singular spectrum analysis combination algorithm. Advances in Structural Engineering 2019, 22(7): 1519-1530, https://doi.org/10.1177/1369433218818921.
12. Li Y, Wang X, Liu Z, Liang X, Si S. The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review. IEEE Access 2018, 6: 66723-66741, https://doi.org/10.1109/ACCESS.2018.2873782.
13. Li Y, Wang X, Si S, Huang S. Entropy based fault classification using the Case Western Reserve University data: A benchmark study. IEEE Transactions on Reliability 2019, 1-14, https://doi.org/10.1109/TR.2019.2896240.
14. Milu Zhang, Tianzhen Wang, Tianhao Tang, Mohamed Benbouzid, Demba Diallo. An imbalance fault detection method based on data normalization and EMD for marine current turbines. ISA Transactions 2017; 68: 302-312, https://doi.org/10.1016/j.isatra.2017.02.011.
15. Saba Adabi, Siavash Ghavami, Mostafa Fatemi, Azra Alizad. Non-Local based denoising framework for in vivo contrast-free ultrasound microvessel imaging. Sensors; 19(2): 245, https://doi.org/10.3390/s19020245.
16. Salman A H , Ahmadi N , Mengko R. Performance Comparison of Denoising Methods for Heart Sound Signal. 2015 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS). IEEE, 2015, https://doi.org/10.1109/ISPACS.2015.7432811.
17. Sikora M, Szczyrba K, Wróbel, Michalak M. Monitoring and maintenance of a gantry based on a wireless system for measurement and analysis of the vibration level. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(2): 341-350, https://doi.org/10.17531/ein.2019.2.19.
18. O.I. Traore, L. Pantera, N. Favretto-Cristini, P. Cristini, S. Viguier-Pla, P. Vieu. Structure analysis and denoising using singular spectrum analysis: Application to acoustic emission signals from nuclear safety experiments. Measurement 2017, 104: 78-88, https://doi.org/10.1016/j.measurement.2017.02.019.
19. Ou D, Tang M, Xue R, Yao H. Hybrid fault diagnosis of railway switches based on the segmentation of monitoring curves. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 514-522, https://doi.org/10.17531/ein.2018.4.2.
20. Te Han, Dongxiang Jiang, Nanfei Wang. The fault feature extraction of rolling bearing based on EMD and difference spectrum of singular value. Shock and Vibration 2016: 5957179, https://doi.org/10.1155/2016/5957179.
21. Tracey B H, Miller E L. Nonlocal means denoising of ECG signals. IEEE Transactions on Bio-medical Engineering 2012; 59(9): 2383, https://doi.org/10.1109/TBME.2012.2208964.
22. Ville D V D, Kocher M. SURE-Based Non-Local Means. IEEE Signal Processing Letters 2009; 16(11): 973-976, https://doi.org/10.1109/LSP.2009.2027669.
23. Wu Ding-hai, Zhang Pei-lin, Yang Wang-can, Qi Yun-guang. Overlappling group thresholding denoising method based on dual-tree complex wavelet packet transform. Journal of Vibration and Shock 2016; 35(10): 162-166.
24. Yang Shaopu, Zhao Zhihong. Improved Wavelet Denoising Using Neighboring Coefficients and Its Application to Machinery Fault Diagnosis. Journal of Mechanical Engineering 2013; 49(17): 137-141, https://doi.org/10.3901/JME.2013.17.137.
25. Yu G, Yin Y, Wang H, et al. Image denoising based on Non-Local means and multi-scale dyadic wavelet transform in IEEE International Conference on Computer Science and Information Technology. IEEE 2010: 333-336.
26. Yongbo Li, Yuantao Yang, Guoyan Li, Minqiang Xu, Wenhu Huang. A fault diagnosis scheme for planetary gearboxes using modified multi-scale symbolic dynamic entropy and mRMR feature selection. Mechanical Systems and Signal Processing 2017, 91: 295-312, https://doi.org/10.1016/j.ymssp.2016.12.040.
27. Zhang Long, Hu Junfeng, Xiong Guoliang. Fault diagnosis of rolling bearings based on weighted nonlocal means algorithm. Journal of Vibration and Shock 2016; 35(19): 156-161, https://doi.org/10.1155/2016/4805383.
28. Zhanxiong Wu, Thomas Potter, Dongnan Wu, Yingchun Zhang. Denoising high angular resolution diffusion imaging data by combining singular value decomposition and non-local means filter. Journal of Neuroscience Methods 2019; 312: 105-113, https://doi.org/10.1016/j.jneumeth.2018.11.020.
29. Zhiwen Liu, Zhengjia He, Wei Guo, Zhangchun Tang. A hybrid fault diagnosis method based on second generation wavelet de-noising and local mean decomposition for rotating machinery. ISA Transactions 2016; 61: 211-220, https://doi.org/10.1016/j.isatra.2015.12.009.

Te HAN
Dongxiang JIANG
Department of Energy and Power Engineering
Tsinghua University
Beijing, 100084, China

E-mail: hant15@mails.tsinghua.edu.cn, jiangdx@tsinghua.edu.cn

Article citation info:
CHERNETS M. Method of calculation of tribotechnical characteristics of the metal-polymer gear, reinforced with glass fiber, taking into account the correction of tooth. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 546–552, http://dx.doi.org/10.17531/ein.2019.4.2.

Myron CHERNETS

METHOD OF CALCULATION OF TRIBOTECHNICAL CHARACTERISTICS OF THE METAL-POLYMER GEAR, REINFORCED WITH GLASS FIBER, TAKING INTO ACCOUNT THE CORRECTION OF TOOTH

METODA OBLICZENIOWA TRIBOTECHNICZNYCH CHARAKTERYSTYK PRZEKŁADNI ZĘBATYCH METAL-POLIMEROWYCH Z POLIAMIDU WZMOCNIONEGO WŁÓKNEM SZKLANYM Z UWZGLĘDNIENIEM KOREKCJI ZĘBÓW

The paper proposes a new method for calculating the service life, wear and contact pressures of metal-polymer gear drives with a correction profile. The effects of height and angular modification in a gear drive made of dispersive glass fibre-reinforced polyamide and steel on its contact and tribocontact parameters are determined. A numerical solution obtained for the gear with height correction has shown that the life of such gear is the longest when the profile correction coefficients $x_1 = -x_2 = 0.1$. *It has been found that the service life of the gear with angular correction is shorter than that of the gear with correction height. The effects of gear tooth height and angular correction on maximum contact pressures and pinion wear are examined and determined.*

Keywords: *method for calculating service life, wear and contact pressures, metal-polymer spur gear drives, dispersive glass fibre-reinforced polyamide, height and angular correction, gear life, contact strength.*

Predstawiono opracowaną nową metodę obliczeniową resursu, zużycia oraz nacisków stykowych przekładni walcowej metal – polimerowej z korekcja uzębienia. Dla przekładni z kołami zębatymi z poliamidu wzmocnionego dyspersyjnym włóknem szklanym i stali zostało przeprowadzone oszacowanie wpływu korekcji technologicznej oraz konstrukcyjnej uzębienia na wskazane parametry kontaktu oraz tribokontaktu. Na podstawie numerycznego rozwiązania zagadnienia dla przypadku korekcji technologicznej zębów kół określono, że największa trwałość przekładni będzie, gdy współczynniki korekcji $x_1 = -x_2 = 0.1$. *Ustalono, że wtedy przy korekcji konstrukcyjnej zębów trwałość przekładni będzie mniejsza nieżeli przy korekcji technologicznej. Został przebadany charakter wpływu korekcji technologicznej oraz konstrukcyjnej zębów na maksymalne naciski stykowe, zużycie zębnika oraz ustalono jego prawidłowości.*

Słowa kluczowe: *metoda obliczeniowa resursu, zużycia oraz nacisków stykowych, przekładnia metal – polimerowa walcowa o zębach prostych, poliamid wzmocniony dyspersyjnym włóknem szklanym, korekcja technologiczna oraz konstrukcyjna, trwałość przekładni, naciski stykowe.*

1. Introduction

Metal-polymer gears are widely used in different areas. For this reason, it is of practical importance to know how to determine their service life or wear with the use of calculation methods. However, neither the literature of the subject nor applied engineering offer such methods, particularly for gears that are made of composite polymers reinforced with glass or carbon particles or fibres to reduce their wear and extend their service life. Only in [23] a simplified calculation method for abrasive wear is applied to determine the life of a spur gear made of dispersive glass or carbon fibre-reinforced polyamide. However, it should be mentioned that abrasive wear is not the most common type of wear for such gears.

The literature of the subject offers methods for calculating the wear of spur gear drives with metal gears [2, 11, 12, 13, 15, 16, 20, 21 et al]. However, these metods are based on the Archard wear equation describing sliding wear, yet this type of wear does not occur in lubricated gear drives, not to mention gear drives with a closed case design. According to results of experimental studies on the wear of dispersive glass or carbon fibre-reinforced polyamide reported in the literature, the dominant type of wear in non-lubricated metal-polymer gear drives is fatigue wear while abrasive wear does not occur here at all. Consequently, it is necessary to devise methods for determining service life and wear of metal-polymer gear drives that take account of the real wear mechanism.

The author and his co-authors have developed methods for calculating contact strength, wear and life of metal gears (straight and skew spur and bevel gears) [4 - 8] that are based on a tribokinetic, mathematical model of material wear caused by sliding friction and take account of their fatigue wear. The proposed calculation method was modified to suit metal-polymer gears, particularly those made of polyamide composites reinforced with dispersive glass or carbon fibres [9]. In [10] the authors report experimental results of their own studies investigating the wear resistance of reinforced polyamide composites used for metal-polymer gear drives to determine their wear resistance in order to be able to implement the proposed numerical solution. Results of both these and other studies reported in the literature point to fatigue wear of polyamide composites reinforced with dispersive glass and carbon fibres.

It should be stressed that the literature of the subject mentions only few experimental studies on polyamide composites used in the design of metal-polymer gear drives [3, 14,17, 18, 19, 26]. For example, the results of numerical modelling of a metal-polymer straight spur gear drive with polyamide 66 gears are presented in [3, 14], where the distribution of load between the gears and its influence on contact and bending stresses is examined and then verified experimentally. The

influence of sliding speed on the frictional force in a tribojoint made of PA6 polymer and S355J2 steel is investigated by disc tests in [26], and the sliding velocities and Hertz contact pressures in the meshing of a metal-polymer spur gear are calculated. The friction coefficients and relative volumetric wear of the teeth of a metal-polymer gear drive from polyamide PA6-Mg, PA6-Na, PA66GF-30, polyoxymethylene POM-C and steel S355 under normal conditions in the air and abrasive media are experimentally investigated in the works [17, 18]. It has been established in [24] that the wear and life of glass and carbon composites largely depend on the volume content of a reinforcing phase. However, there are no other studies of this nature in the literature. More information about the contact and bending strengths of metal-polymer gears determined by well-known methods is given in [1, 23, 25, 27] dentition correction is used.

Tooth profiles in metal and metal-polymer gear drives are often corrected, which leads to reduced contact pressures and wear of the engaged teeth and increased gear life. The literature of the subject, however, contains no mention of numerical or experimental studies investigating the effect of gear tooth profile correction on the load capacity and tribotechnical parameters of metal-polymer gears. In light of the above, the author of this paper – using his own methods – has undertaken such study for a spur gear drive made of steel and dispersive glass fibre-reinforced polyamide with height and angular correction of the gear teeth, and obtained results are given below.

2. Method for solving the problem

The method for assessing the wear of metal-polymer gear drives is based on a tribokinetic mathematical model of sliding abrasive wear [4, 6] shown below. According to this model, the wear of engaged teeth is described with a system of linear differential equations:

$$\frac{1}{v}\frac{dh_k}{dt} = \Phi_k^{-1}(\tau),\ \ k = 1;2\,, \tag{1}$$

where $\tau = fp$.

Experimental values of the wear-resistance function $\Phi(\tau)$ of the materials are approximated by the relation:

$$\Phi_k(\tau) = C_k\left(\frac{\tau_S}{\tau}\right)^{m_k}, \tag{2}$$

where $\tau_S = R_{0,2}/2$; $R_{0,2} = 0{,}7R_m$ (steel), $\tau_S = R_m/2$ (filled polymer composites).

The wear-resistance function $\Phi_i(\tau_i)$ of the teeth materials is determined in the following way:

$$\Phi_i(\tau_i) = L/h_i\,.$$

Taking into account relation (2), after the separation of variables and system integration (1) on condition that $\tau = fp = \text{const}$, the following will arise:

$$t_k = \frac{C_k}{v}\left(\frac{\tau_S}{\tau}\right)^{m_k} h_k\,. \tag{3}$$

Then, the function of linear wear of the teeth at any point j of the working surface over a period t_j of their interaction:

$$h_k = \frac{vt_k}{C_k}\left(\frac{\tau}{\tau_S}\right)^{m_k}. \tag{4}$$

The linear wear of the gear teeth h'_{kj} at any point j of the profile in the tooth engagement time t'_j is determined using the following formula [4]:

$$h'_{kj} = \frac{v_j t'_j (fp_{j\max})^{m_k}}{C_k \tau_S{}^{m_k}} \tag{5}$$

where $j = 0,1,2,3,...,\ s,\ t'_j = 2b_j/v_0$, $v_0 = \omega_1 r_1 \sin\alpha_t$.

Tooth wear causes an increase in the curvature radii of tooth profiles, which leads to a decrease in the initial maximum contact pressures $p_{j\max}$, and the contact area width $2b_j$ at every j - th point of contact are calculated in accordance with the Hertz equations:

$$p_{j\max} = 0.564\sqrt{N'\theta/\rho_j}\ ,\ 2b_j = 2.256\sqrt{\theta N'\rho_j}\ , \tag{6}$$

where $N' = N/l_{\min}w$; $N = 9550P/r_1 n_1 \cos\alpha_t$; respectively, for the spur gear with the pinion width $l_{\min} = b_W$; $\theta = \left(1-\mu_1^2\right)/E_1 + \left(1-\mu_2^2\right)/E_2$.

The reduced radius of curvature of the involute spur gear is:

$$\rho_j = \frac{\rho_{1j}\rho_{2j}}{\rho_{1j}+\rho_{2j}}\,. \tag{7}$$

The formulas for calculating the radii of curvature for the modified pinion and gear profiles of the spur gear at j - th point of contact are [5]:

$$\rho_{1j} = \frac{\rho_{t1j}}{\cos\beta_b}\ ,\ \rho_{2j} = \frac{\rho_{t2j}}{\cos\beta_b}\ , \tag{8}$$

where for the bevel gear:

$$\beta_b = \arctan\left(\tan\beta\cos\alpha_t\right),\ \ \alpha_t = \arctan\left(\frac{\tan\alpha}{\cos\beta}\right),$$

$$\rho_{t1j} = r_{b1}\tan\alpha_{t1j}\ ,\ \ \rho_{t2j} = r_2\sqrt{\left(r_{2j}/r_2\right)^2 - \cos^2\alpha_t}\ ,$$

$$\alpha_{t1j} = \arctan\left(\tan\alpha_{t10} + j\Delta\varphi\right),\ \alpha_{t1s} = \arctan\sqrt{\left(r_{1s}/r_1\right)^2 - \cos^2\alpha_t}\ ,$$

$$\alpha_{t2j} = \arccos\left[\left(r_2/r_{2j}\right)\cos\alpha_t\right],$$

$$r_{b1} = r_1\cos\alpha_t\ ,\ r_1 = mz_1/2\cos\beta,\ \ r_{b2} = r_2\cos\alpha_t\ ,\ r_2 = mz_2/2\cos\beta\ ,$$

$$\tan\alpha_{t10} = \left(1+u\right)\tan\alpha_t - \frac{u}{\cos\alpha_t}\sqrt{\left(r_{20}/r_2\right)^2 - \cos^2\alpha_t}\ ;\ r_{a2} = r_2 + m\ ,$$

$$r_{20} = r_{a2} - r\ ,\ \ r = 0.2m\ ,$$

$$r_{2j} = \sqrt{a_W^2 + r_{1j}^2 - 2a_W r_{1j}\cos\left(\alpha_t - \alpha_{t1j}\right)}\ ,\ r_{1j} = r_1\cos\alpha_t/\cos\alpha_{t1j}\ ,$$

$$a_W = \left(z_1 + z_2\right)m/2\cos\beta\,.$$

The minimum length of the line of contact is:

$$l_{\min} = \frac{b_W \varepsilon_\alpha}{\cos\beta_b}\left[1 - \frac{(1-n_\alpha)(1-n_\beta)}{\varepsilon_\alpha \varepsilon_\beta}\right] \text{ at } n_\alpha + n_\beta \rangle 1,$$

$$l_{\min} = \frac{b_W \varepsilon_\alpha}{\cos\beta_b}\left[1 - \frac{n_\alpha n_\beta}{\varepsilon_\alpha \varepsilon_\beta}\right] \text{ at } n_\alpha + n_\beta \le 1, \quad (9)$$

where: $\varepsilon_\alpha = \frac{t_1 + t_2}{t_z}$, $\varepsilon_\beta = \frac{b_W \sin\beta}{\pi m}$, $\varepsilon_\gamma = \varepsilon_\alpha + \varepsilon_\beta$, $t_1 = \frac{e_1}{\omega_1 r_{b1}}$, $t_2 = \frac{e_2}{\omega_1 r_{b1}}$, $t_z = \frac{2\pi}{z_1 \omega_1}$, $e_1 = \sqrt{r_{1s}^2 - r_{b1}^2} - r_1 \sin\alpha_t$, $e_2 = \sqrt{r_{20}^2 - r_{b2}^2} - r_2 \sin\alpha_t$, $r_{1s} = r_{a1} - r$, $r_{a1} = r_1 + m$.

In Fig. 1 presents spur gear engagement and transmission parameters. The sliding velocity of the engaged teeth is calculated as [5]:

$$v_j = \omega_1 r_{b1}\left(\tan\alpha_{t1j} - \tan\alpha_{t2j}\right). \quad (10)$$

In a simplified case, at constant output conditions, i.e., when the initial contact pressures $p_{j\max} = const$, the gear life t_* for a given acceptable tooth wear h_{k*} is calculated as:

$$t_* = h_{k*} / \bar{h}_{kj}, \quad (11)$$

where $\bar{h}_{kj} = 60 n_k h'_{kj}$.

In some of the above formulas one should consider the modified engagement parameters.

For the gears with height correction:

The addendum radii of the gear:

$$r_{a1} = r_1 + (1 + x_1)m, \quad r_{a2} = r_2 + (1 + x_2)m, \quad (12)$$

where $x_1 = -x_2$.

The remaining parameters of the gear are the same as those of the gear without profile correction.

For the gears with angular correction:

Here $x_1 \ne x_2$, and the total profile shift coefficient $x_\Sigma = x_1 + x_2$. The working (real) distance between the axes is:

$$a_{wk} = r_{w1} + r_{w2} \rangle a_w. \quad (13)$$

The corrected profile pressure angle α_w depends on the real distance between the axes of the meshing gears and is higher (when $a_{wk} > a_w$) than the apparent pressure angle α_t. If the real distance between the axes is known, then:

$$\alpha_w = \arccos\frac{a_w}{a_{wk}}\cos\alpha_t. \quad (14)$$

The pitch radii of the pinion and gear teeth:

$$r_{w1} = r_1 \frac{\cos\alpha_t}{\cos\alpha_w}, \quad r_{w2} = r_2 \frac{\cos\alpha_t}{\cos\alpha_w}. \quad (15)$$

The addendum radii of the gear teeth:

$$r_{a1} = r_1 + (1 + x_1 - K)m, \quad r_{a2} = r_2 + (1 + x_2 - K)m, \quad K = \frac{a_w - a_{wk}}{m} + x_\Sigma. \quad (16)$$

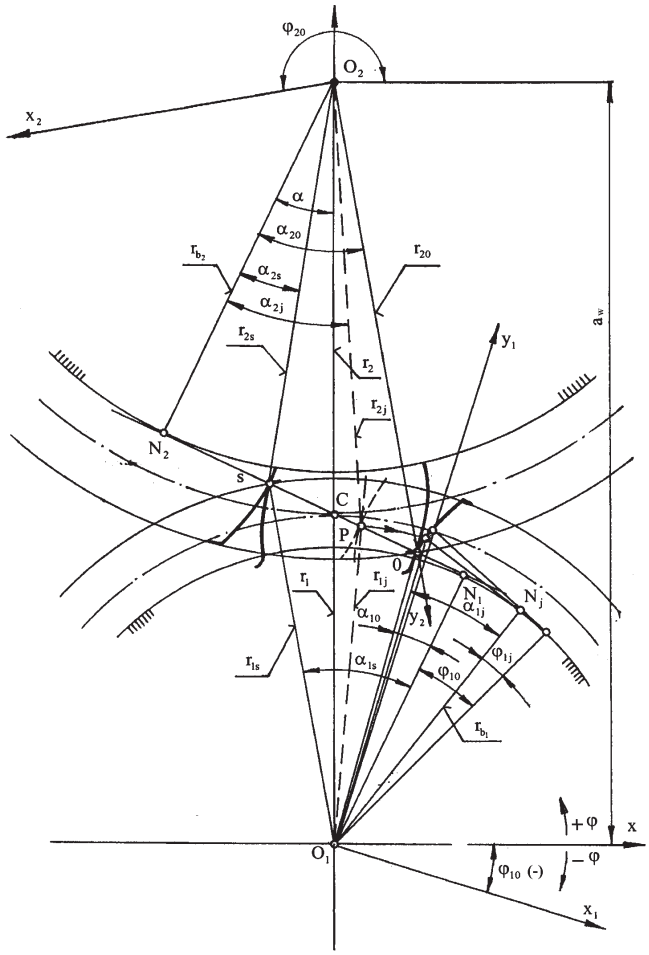


Fig. 1. Parameters of gear and engagement: N_1N_2 - line of engagement; $0, s$ - respectively, the points of entry of the teeth in the engagement teeth and exit; C – engagement means

Formulas in which the above profile correction parameters should be taken into account are as follows:

$$N = 9550 P K_g / r_{w1} n_1 \cos\alpha_w,$$

$$tg\alpha_{t10} = (1+u)tg\alpha_w - \frac{u}{\cos\alpha_w}\sqrt{(r_{20}/r_{w2})^2 - \cos^2\alpha_w},$$

$$\alpha_{t1s} = arctg\sqrt{(r_{1s}/r_{w1})^2 - \cos^2\alpha_w} \quad \rho_{t2j} = r_{w2}\sqrt{(r_{2j}/r_{w2})^2 - \cos^2\alpha_w},$$

$$r_{2j} = \sqrt{a_w^2 + r_{1j}^2 - 2a_w r_{1j}\cos(\alpha_w - \alpha_{t1j})}, \quad r_{1j} = r_{w1}\cos\alpha_w / \cos\alpha_{t1j},$$

$$tg\alpha_{t2s} = (1 + u^{-1})tg\alpha_w - \frac{1}{u\cos\alpha_w}\sqrt{(r_{1s}/r_{w1})^2 - \cos^2\alpha_w},$$

$$\alpha_{t2j} = \arccos\left[(r_{w2}/r_{2j})\cos\alpha_w\right], e_1 = \sqrt{r_{1s}^2 - r_{b1}^2} - r_{w1}\sin\alpha_w,$$

$$e_2 = \sqrt{r_{20}^2 - r_{b2}^2} - r_{w2}\sin\alpha_w.$$

The angles of transition from double-pair engagement ($\Delta\varphi_{1F_2}$) to single-pair engagement and, again, to double-pair engagement ($\Delta\varphi_{1F_1}$) in the spur gear with a corrected profile are determined in the following way [6, 7]:

$$\Delta\varphi_{1F_2} = \varphi_{10} - \varphi_{1F_2},\ \Delta\varphi_{1F_1} = \varphi_{10} + \varphi_{1F_1}; \tag{17}$$

where $\varphi_{1F_2} = \tan\alpha_{F_2} - \tan\alpha_t$, $\varphi_{1F_1} = \tan\alpha_{F_1} - \tan\alpha_t$, $\phi_{10} = \tan\alpha_{t10} - \tan\alpha_w$;

$$\tan\alpha_{F_2} = \frac{r_1 \sin\alpha_t - (p_b - e_1) + 0.5 n_\beta p_b}{r_1 \cos\alpha},\ \tan\alpha_{F_1} = \frac{r_1 \sin\alpha_t - (p_b - e_2) - 0.5 n_\beta p_b}{r_1 \cos\alpha}. \tag{18}$$

In the spur gear e_β, $n_\beta = 0$ and $l_{\min} = b_W$, $p_b = \pi m \cos\alpha_w$.

The angle $\Delta\varphi_{1E}$ describing the moment of pair engagement exit is:

$$\Delta\varphi_{1E} = \varphi_{10} + \varphi_{1E}, \tag{19}$$

where $\varphi_{1E} = \tan\alpha_E - \tan\alpha_t$, $\alpha_E = \arccos(r_{b1} / r_{1s})$; $p_b = \pi m \cos\alpha_t / \cos\beta$.

For triple-double-triple pair engagement (helical gears):

$$\tan\alpha_{F_2} = \frac{r_1 \sin\alpha_t - (p_b - e_1) + 0.5 \tilde{n}_\beta p_b}{r_1 \cos\alpha},\ \tan\alpha_{F_1} = \frac{r_1 \sin\alpha_t - (p_b - e_2) - 0.5 \tilde{n}_\beta p_b}{r_1 \cos\alpha}; \tag{20}$$

$$\tilde{n}_\beta = \begin{cases} n_\beta & at\ n_\alpha + n_\beta \rangle 1, \\ 1 - n_\beta & at\ n_\alpha + n_\beta \le 1 \end{cases}.$$

3. Numerical solution

The calculations were performed for a metal-polymer spur gear drive with a glass fibre-reinforced polyamide pinion and a steel gear after the application of height and angular correction of the gear teeth. Initial data applied in the calculations were as follows: T_{nom} = 4000 Nmm, n_1 = 1000 rpm; K_g =1.2; β = 0; m = 4 mm, u = 3, z_1 = 20, z_2 = 60, b_W = 50 mm, f = 0.3; h_{k*} = 0.5 mm.

The applied profile correction coefficients were: a) height correction: $x_1 = -x_2$ = 0; 0.1; 0.2; 0.3; 0.4; a_W = 160 mm; b) angular correction: $x_1 = 0,\ x_2 \rangle 0;\ x_1 \rangle 0,\ x_2 = 0;\ x_1 = 0,\ x_2 \langle 0;\ a_{Wk}$ = 161 mm.

The gear materials:

1) steel S45 in the state of delivery, $E = 2.1 \cdot 10^5$ MPa, μ = 0.3; $C = 10^9$, m = 2 [10];
2) glass fibre-reinforced polyamide composite (30% vol.) PA6-LT-GF30-1, E_G = 3,90 GPa, μ_G = 0.42; $C_G = 1.2 \cdot 10^6$, m_G = 1.9 [10]; $\tau_S^{(G)}$ = 52 MPa.

In the above wear resistance characteristics of the gear materials, particularly of the polyamide composite with glass fibres PA6-LT-SW30-1 [10] for the unit friction force $\tau = fp$, the characteristic m_G = 1.9 indicates a nearly quadratic dependence between the experimental wear function (Eq. (2)) and the contact pressures $p = \tau / f$. This means that we are dealing here with fatigue wear rather than abrasive wear, because in the latter case m_G = 1. The wear of the steel gear is three times lower than that of the polyamide gear.

The results are given in Figs. 2 - 7. Fig. 1 illustrates the relationship between the minimal gear life $t_{\min}$ at the contact point on the gear tooth profile where the maximum allowable wear occurs faster.

The results demonstrate that the application of height correction when $x_1 = -x_2 = 0.1$ increases the gear life $t_{\min}$ by 1.2 times compared to the life of the gear without profile correction. On the other hand, when $x_1 = -x_2 > 0.18$, the gear life decreases compared to both the gear without profile modification and the above-mentioned gear with angular correction (dashed-dotted line).

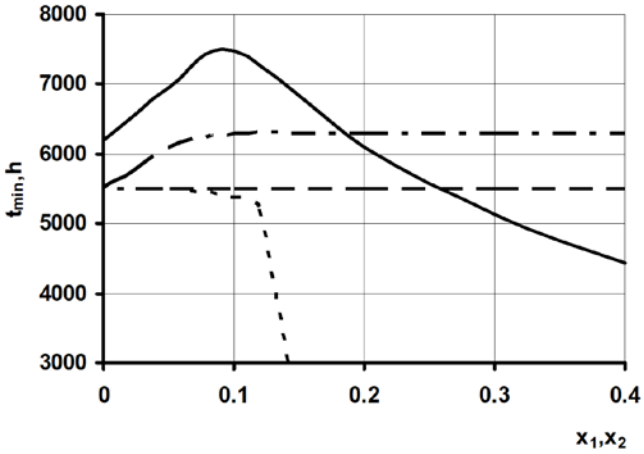


Fig. 2. Minimal life $t_{\min}$ of a gear with profile engagement: solid line: $x_1 = -x_2$ = 0 ... 0.4; dashed-dotted line: $x_1 = 0$ ... 0.4; $x_2 = 0$; dashed line: $x_1 = 0$, x_2 = 0 ... 0.4; dotted line: $x_1 = 0$, $x_2 = 0$ ... -0.1462

Among the analysed cases of a gear drive with angular correction where a_{wk} = 161 mm, the life of the gear drive somewhat increases when $x_1 = 0$ ... 0.1; $x_2 = 0$. Nevertheless, the gear life is lower by 1.19 times than that of the gear with height correction. As regards two other cases of the gear with angular correction, the application of angular correction either has no effect on gear life ($x_1 = 0$, x_2 = 0 … 0.4) or leads to its sudden decrease ($x_1 = 0$, $x_2 = 0$ ... -0.1462).

The minimal life $t_{\min}$ of the gear drive with a steel pinion and a polyamide composite gear was determined. It has been found that the minimal life of the gear increases in direct proportion to the gear ratio u, i.e., by three times in this particular case.

An important parameter describing the meshing conditions is the overlap factor ε_α. Fig. 3 illustrates the effect of different types of profile correction on the overlap factor in double-single-double pair engagement.

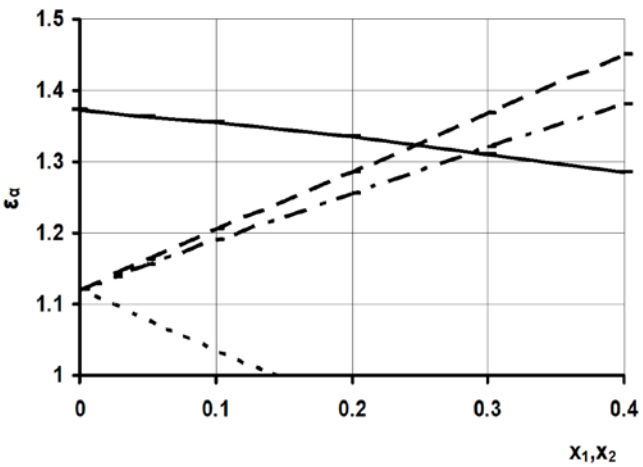


Fig. 3. Overlap factor vs. double-single-double pair engagement

The overlap factor ε_α of the gear with height correction for the optimal modification coefficients $x_1 = -x_2 = 0.1$ is definitely higher than in the case of the gear with angular correction when $x_1 = 0.1$; $x_2 = 0$

or $x_1 = 0;\ x_2 = 0.1$, which has a positive effect on service life of the gear drive (Fig. 2).

Fig. 4. illustrates variations in the maximum contact pressures for the meshing gears when height correction is applied as the optimal type of profile modification. Their values decrease between the entry and exit of pair engagement due to an increase in the reduced radius of curvature and the changes in gear pair engagement (double–single–double pair engagement).

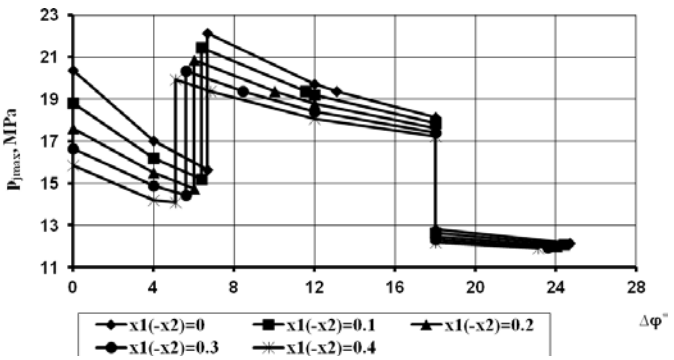


Fig. 4. Effect of height correction on maximum contact pressures in a gear meshing cycle

The results reveal a significant influence of height correction on the value of $p_{j\max}$ in a gear meshing cycle. The greatest changes can be observed at the entry of double-pair engagement when $\Delta\varphi = 0$ (left side of the figure) and at the entry of single-pair engagement in the central region. The highest contact pressures occurring at the entry of single-pair engagement are higher than in the centre of the meshing gears (19.36 MPa – all markers are on the same level). Therefore, it can be observed that the application of profile correction leads to decreasing the level of $p_{\max}$ in the entire tested range of addendum correction coefficients $x_1 = -x_2 \rangle 0$, whereas at $x_1 = -x_2 \approx 0.2$ the life of the modified gear drive is lower than that of the gear without profile correction (Fig. 2). The selection of profile correction coefficients should depend on contact strength of the gear teeth or gear life, or both.

The investigation of the effect of angular correction when $x_1 = 0 \ldots 0.4;\ x_2 = 0$ demonstrates that the gear life increases while the contact pressures $p_{j\max}$ decrease. Results of these calculations are given in Fig. 5.

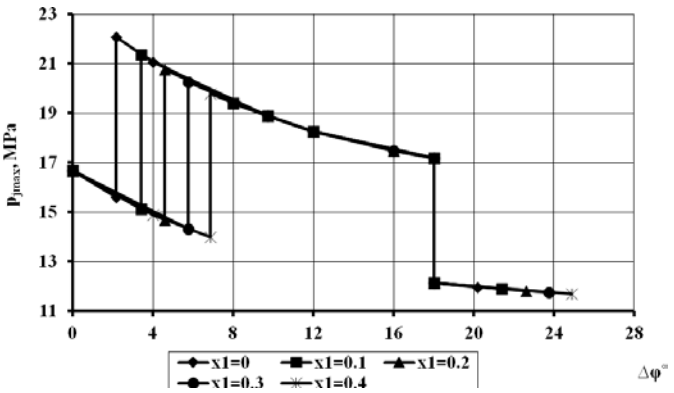


Fig. 5. Effect of angular correction on maximum contact pressures in a gear meshing cyclewhen $x_1 = 0 \ldots 0.4;\ x_2 = 0$

The correction of the pinion teeth only affects the maximum contact pressures $p_{j\max}$ at the exit of double-pair engagement and at the entry of single-pair engagement. Contrary to height correction, increasing the coefficient x_1 leads to increasing the double-pair engagement area, which results in a slight reduction in the contact pressures at this point without any changes in the service life of the gear drive (Fig. 2), which remains unchanged starting from $x_1 \approx 0.1$. Hence, this type of angular correction can be applied to reduce maximum contact pressure when the service life of the gear with height correction of $x_1 = -x_2 > 0.18$ decreases (Fig. 2).

In effect of the application of angular correction when $x_1 = 0, x_2 = 0 \ldots 0.4$, both the highest contact pressures occurring at the entry of single-pair engagement and the gear life remain unchanged (Fig. 6). As in the previous case angular correction increasing the coefficient x_2 leads to increasing the double-pair engagement area, but this does not increase the durability of the transmission. Instead, the negative consequence of the correction of the teeth of the metal wheel is the increase in contact pressures at the entrance to the double-pair engagement. Therefore, this case of angular correction engagement is not appropriate.

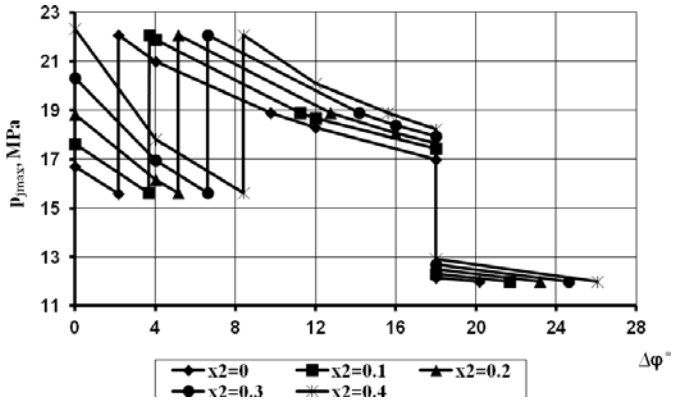


Fig. 6. Effect of angular correction on maximum contact pressures in a gear meshing cycle when $x_1 = 0, x_2 = 0 \ldots 0.4$

Fig. 7.shows the linear wear of the pinion teeth when height correction is applied.

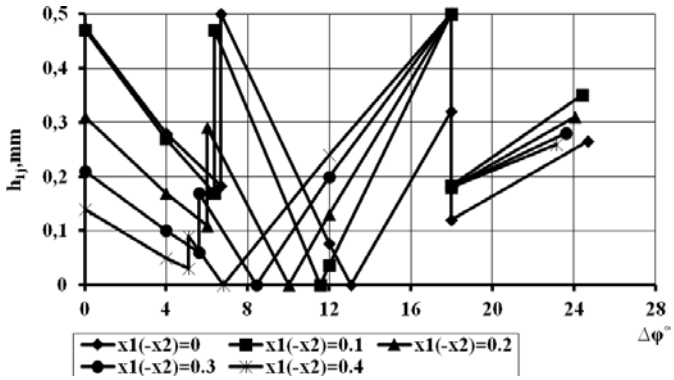


Fig. 7. Linear wear of the pinion teeth in a gear meshing cycle

The results of wear obtained for the non-corrected gear drive and the modified gear with the correction coefficients $x_1 = -x_2 = 0.1$ are similar. The latter case is optimal, and the wear is almost identical at three points of the tooth profile: at the entry of both double- and single-pair engagement as well as at the exit of single-pair engagement (acceptable wear), which undoubtedly leads to a longer service life of the gear drive (Fig. 2). In other cases, the wear at these three particular points varies, while the acceptable wear is reached at the exit of single-pair engagement.

Given the results of maximum contact pressures, service life and wear of metal-polymer gear drives obtained with the developed cal-

culation method, it is recommended to apply height correction with relatively small modification coefficients ($x_1 = -x_2 \approx 0.15$) in order to increase the load carrying capacity and service life of a gear drive.

5. Conclusions

1. The proposed calculation method is employed to thoroughly investigate metal-polymer gear drives with gears made of dispersive glass fibre-reinforced polyamide composite in order to assess their strength properties and tribotechnical parameters as well as to verify whether profile correction is necessary, which is of great practical important at the design stage.
2. The results of problem solution demonstrate that the application of height correction in a limited range of variation of addendum correction coefficients leads to an increase in the gear life compared to the gear without profile correction (Fig. 2).
3. When $x_1 = -x_2 = 0.1$, the gear life increases by 12% and is the highest possible for the gear with angular correction.
4. The maximum contact pressures in the engagement cycle vary significantly, depending on the engagement area and the pinion rotation angle. They reach the greatest value at the entry of single-pair engagement (Fig. 4 – height correction, Fig. 5 – angular correction).
5. In the case of the gear drive with height correction when $x_1 = -x_2 = 0.1$, the wear of the pinion teeth is almost identical at three points on the tooth profile – at the entry of double- and single-pair engagement as well as at the exit of single-pair engagement (Fig. 7). In other cases, when $x_1 = -x_2 > 0.1$, the wear at these three points of engagement varies and the acceptable wear is reached at the exit of single-pair engagement.
6. It is rational to apply height correction in compliance with the criteria of gear life and contact pressures when $x_1 = -x_2 < 0.18$, while angular correction should be applied according to these criteria when $x_1 > 0.18$, $x_2 = 0$.
7. In metal-polymer gears, heat can occur due to the action of frictional forces in the engagement. In particular, such a phenomenon is intensified with an increase in power or load. This can increase the wear of the wheel teeth. Therefore, this should be taken into account when operating this kind of gear.

References

1. Bharat G, Abhishek C, and Gautam V V. Contact stress analysis of spur gear. International Journal of Engineering Research and Technology 2012; 1: 1-7.
2. Brauer J, Andersson S. Simulation of wear in gears with flank interference - a mixed FE and analytical approach. Wear 2003; 254: 1216-1232, https://doi.org/10.1016/S0043-1648(03)00338-7.
3. Cathelin J, Letzelter E, Guingand M, De Vaujany J P, Chazeau L. Experimental and Numerical Study a Loaded Cylindrical PA66 Gear. Journal of Mechanical Design 2013; 135: 89-98, https://doi.org/10.1115/1.4023634.
4. Chernets M V, Kelbinski J, Jarema R Ja. Generalized method for the evaluation of cylindrical involute gears. Materials Science 2011; 1: 45-51, https://doi.org/10.1007/s11003-011-9366-9.
5. Chernets M ., Yarema R Y, Chernets J M. A method for the evaluation of the influence of correction and wear of the teeth of a cylindrical gear on its durability and strength. Part 1. Service live and wear. Materials Science 2012; 3: 289-300, https://doi.org/10.1007/s11003-012-9505-y.
6. Chernets M V, Chernets J M. Evaluation of the strength, wear, and durability of a corrected cylindrical involute gearing, with due regard for the engagement conditions. Journal of Friction and Wear 2016; 37 (1): 71-77, https://doi.org/10.3103/S1068366616010050.
7. Chernets M V, Chernets Y M. A technique for calculating tribotechnical characteristics of tractive cylindrical gear of VL - 10 locomotive. Journal of Friction and Wear 2017; 37 (6): 566-572, https://doi.org/10.3103/S1068366616060040.
8. Chernets M, Chernets J. The simulation of influence of engagement conditions and technological teeth correction on contact strength, wear and durability of cylindrical spur gear of electric locomotive. Proc. J. Mech. E. Part J: Journal of Engineering Tribology 2017; 231 (1): 57-62, https://doi.org/10.1177/1350650116645024.
9. Chernets M, Shil'ko S, Pashechko M. Study of wear resistance of reinforced polyamide composites for metal-polymer gear drives. Tribologia 2018; 3: 19 - 23, https://doi.org/10.5604/01.3001.0012.7003.
10. Chernets M V, Shil'ko S V, Pashechko M I, and Barshch M. Wear resistance of glass- and carbon-filled polyamide composites for metal-polymer gears. Journal of Friction and Wear 2018; 39 (5): 361 - 364, https://doi.org/10.3103/S1068366618050069.
11. Drozdov Y. To the development of calculation methods on friction wear and modeling. Wear resistance. Science, Moscow: Science, 1975: 120-135.
12. Flodin A, Andersson S. Wear simulation of spur gears. Tribotest J. 1999; 3(5): 225-250, https://doi.org/10.1002/tt.3020050303.
13. Flodin A, Andersson S. A simplified model for wear prediction in helical gears. Wear 2001; 249 (3-4): 285-292, https://doi.org/10.1016/S0043-1648(01)00556-7.
14. Hooke C J, Kukureka S N, Liao P, Rao M, Chen Y K. The Friction and Wear of Polymers in Non-Conformal Contacts. Wear 1996; 200: 83-94, https://doi.org/10.1016/S0043-1648(96)07270-5.
15. Grib V. Solution of tribotechnical tasks with numerous methods. Moscow: Science, 1982.
16. Kahraman A, Bajpai P, Anderson N E. Influence of tooth profile deviations on helical gear wear. J. Mech. Des. 2005; 127 (4): 656-663, https://doi.org/10.1115/1.1899688
17. Kalacska G, et al. Friction and Wear of Engineering Polymer Gears. Proceedings of WTC2005 World Tribology Congress III, Sept. 12-16, 2005, Washington, https://doi.org/10.1115/WTC2005-63961.
18. Keresztes R, Kalacska G. Friction of Polymer/Steel Gear Pairs. Plastics and Rubber 2008; 45: 236-242.
19. Kindrachuk M V, Volchenko A I, Volchenko D A, Zhuravlev D Y, Chufus V M. Elektrodynamics of the Termal Contact Friction Interaction in Metal-Polymer Friction Couples. Material Science 2018; 54 (1): 69 - 77, https://doi.org/10.1007/s11003-018-0159-2.
20. Kolivand M, Kahraman A. An ease-off based method for loaded tooth contact analysis of hypoid gears having local and global surface deviations. J. Mech. Des. 2010; 132 (7): 0710041-0710048, https://doi.org/10.1115/1.4001722.
21. Pasta A, Mariotti Virzi G. Finite element method analysis of a spur gear with a corrected profile. J. Strain Analysis 2007; 42: 281-292, https://doi.org/10.1243/03093247JSA284.
22. Pronikov A. Reliability of machines. Moscow: Mashinostroenie, 1978.

23. Sajad H D, Vivek A, Mohammad J K, Arunish M. Investigation of bending stress on a spur gear tooth at design stage by finite element modelling. International Journal on Mechanical Engineering and Robotics 2015; 3: 13-18.
24. Shil'ko S V, Starzhinskii V E. Prediction of Wear Resistance of Gearing with Wheels Made of Reinforced Composites. Journal of Friction and Wear 1993; 14 (3): 7-13.
25. Shil'ko S V, Starzhinsky V E, Petrokovets E M, Chernous D A. Two-Level Calculation Method for Tribojoints Made of Disperse-Reinforced Composites: Part 1. Journal of Friction and Wear 2013; 34 (1): 65-69, https://doi.org/10.3103/S1068366613010133.
26. Sukumaran J, et al. Modelling Gear Contact with Twin-Disc Setup. Tribology International 2012; 49: 1-7, https://doi.org/10.1016/j.triboint.2011.12.007.
27. Sushil Kumar T, Upendra Kumar J. Stress analysis of mating involute spur gear teeth. International Journal of Engineering Research and Technology 2012; 1: 1-12.

Notations

a_W – centre distance of the gear pair

a_{wk} – real distance between the axes

b_W – width of the pinion

C_k, m_k – wear resistance characteristics of the gear materials

E, μ – Young's modulus and the Poisson's ratio of the gear tooth materials, respectively

f – sliding friction factor

h – linear wear of the material of the tribosystem elements

h_i – linear wear of the material samples

$\overline{h}_{kj}$ – linear wear of the gear teeth per hour

h'_{kj} – single linear wear of the teeth at any j-th point of the profile

h_{k*} – acceptable wear of the composite gear

$i = 1, 2, 3, \ldots$ loading ratios

j – contact points on the active face of the teeth

$j = 0$, $j = s$ – first and last point of tooth engagement, respectively

K_g – dynamic factor

$k = 1; 2$ – numbers of the gears (1 – pinion, 2 – gear)

K – addendum modification coefficient

L – friction length

$l_{\min}$ – minimum length of the contact line

m – engagement module

$n = n_k = 1, 2, 3, \ldots$ – number of gear revolutions

n_1 – number of pinion revolution

n_α, n_β – fractional parts of the coefficients $\varepsilon_\alpha, \varepsilon_\beta$

N – force acting in the engagement

p – maximum tribocontact pressure

$p_{j\max}$ – maximum tribocontact pressure (at tooth wear) at j - th point of contact

p_b – pitch of the teeth

P – power on the drive shaft (pinion)

r – radius of the gear tooth fillet

r_1, r_2 – pitch circle radii of the pinion and gear, respectively

r_{b1}, r_{b2} – base circle radii of the pinion and gear, respectively

r_{a1}, r_{a2} – addendum circle radii of the pinion and gear, respectively

r_{w1} – rolling radius of the pinion

$R_{0,2}$ – conventional yield strength of the material in tension

R_m – immediate tensile strength of the material

t – time of wear

t'_j – time of wear of the teeth in the course of motion of j -th point of their contact along the tooth by the width of contact zone $2b_j$

$t_{*\min}$ – minimal gear life $t_{*\min}$ for the highest tooth wear $\overline{h}_{kj}$

T_{nom} – rated torque

u – gear ratio

$v_j = v$ – sliding velocity at j-th point of the tooth profile

v_0 – velocity of the contact point along the tooth profile

w – number of the engaged tooth pairs

x_1, x_2 – addendum modification coefficients

z_1, z_2 – number of teeth of the pinion and the gear, respectively

α_t – apparent pressure angle

α_w – pressure angle of the modified profile

α = 20° – pressure angle of the engaged teeth

α_{t10} – angle of the first point on the contact line

α_{t1s} – angle describing the location of the last point of engagement of the pinion teeth on the contact line

$\alpha_{t20}, \alpha_{t2s}$ – angles describing the location of the first and last points of engagement of the gear teeth on the contact line

β – tooth pitch angle

$\varepsilon_\alpha, \varepsilon_\beta$ – coefficients describing the top and step-by-step overlap of the gear

$\Phi(\tau)$ – function of wear resistance of the gear drive materials

$\Delta\varphi$ – selected angle of rotation of the teeth of the pinion from the point of initial contact (point 0) to point 1, and so on

ρ_j – reduced radius of curvature of the gear profile changeable due to wear, in a normal section

ρ_{1j}, ρ_j – changeable curvature radii of the pinion and gear tooth profiles, respectively

ρ_{1j}, ρ_{2j} – curvature radii of the unworn tooth flank profiles of the pinion and the gear, respectively

τ – specific friction force

τ_S – shear strength of the material

ω_1 – angular velocity of the pinion

Myron CHERNETS
Department of Mechanical Engineering
Lublin University of Technology
ul. Nadbystrzycka 36, 20-816 Lublin, Poland

E-mail: m.czerniec@pollub.pl

Article citation info:
PAŚ J, KLIMCZAK T. Selected issues of the reliability and operational assessment of a fire alarm system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 553–561, http://dx.doi.org/10.17531/ein.2019.4.3.

Jacek PAŚ
Tomasz KLIMCZAK

SELECTED ISSUES OF THE RELIABILITY AND OPERATIONAL ASSESSMENT OF A FIRE ALARM SYSTEM

WYBRANE ZAGADNIENIA OCENY NIEZAWODNOŚCIOWO-EKSPLOATACYJNEJ SYSTEMÓW SYGNALIZACJI POŻARU*

The article discusses an analysis of the operational and reliability issues, which regards selected fire alarm systems (FAS) exhibiting different functional structures. These systems are operated in a vast transport area, within a specific environment. We can distinguish three basic structures of these systems – focused, dispersed and mixed. A given system functional structure, utilized within a facility (a given area) is a function depending on the configuration, internal connections of elements and devices, and a developed fire scenario. The application of a given system structure for fire protection also depends on the legislation determining the approval of a facility (area) for use. The process of executing a scenario in the event of a fire is ensured by an algorithm implemented in the alarm central unit and other elements of the system. The implementation of all the system requirements specified within a given procedure algorithm depends on, e.g., an appropriate reliability structure and environmental conditions. The article analyses the operational process of selected FAS, which are operated within a vast transport area. It discusses the actual results of the operational process tests, e.g., repair and damage durations. Next, operational relationship graph, taking into account the conducted operational test, was developed. This enabled the determination of relationships that allow to specify the operating and reliability parameters in terms of a FAS staying in the states distinguished for the research. The FAS test methodology presented in the article, owing to meeting specific performance requirements, can be used in the course of developing a fire scenario and designing systems, taking into account various available technical solutions.

***Keywords**: operation, reliability, fire alarm systems.*

W artykule przeprowadzono analizę problemów eksploatacyjnych i niezawodnościowych, która dotyczy wybranych systemów sygnalizacji pożaru (SSP) o różnej strukturze funkcjonalnej. Systemy te są użytkowane na rozległym obszarze transportowym, w określonym środowisku. Można wyróżnić trzy podstawowe struktury tych systemów - skupiona, rozproszona i mieszana. Dany rodzaj struktury funkcjonalnej systemu, który jest użytkowany w obiekcie (na danym obszarze) jest funkcją zależną od konfiguracji, wewnętrznych połączeń elementów i urządzeń oraz opracowanego scenariusza postępowania na wypadek pożaru. Zastosowanie danej struktury systemu do ochrony pożarowej zależy także od przepisów prawnych warunkujących dopuszczenie danego obiektu (obszaru) do użytkowania. Proces realizacji scenariusza w czasie pożaru jest gwarantowany przez algorytm zaimplementowany w centrali alarmowej oraz innych elementach systemu. Realizacja wszystkich wymagań wobec systemu określonych w danym algorytmie postępowania uwarunkowana jest np. odpowiednią strukturą niezawodnościową i warunkami środowiskowymi. W artykule przedstawiono analizę procesu eksploatacji wybranych SSP, które są użytkowane na obszarze transportowym. Zaprezentowano rzeczywiste wyniki badań procesu eksploatacji, np. czasy trwania naprawy oraz uszkodzenia. Następnie opracowano graf relacji eksploatacyjnych z uwzględnieniem przeprowadzonych badań eksploatacyjnych. Umożliwiło to wyznaczenie zależności pozwalających na określenie parametrów eksploatacyjnych i niezawodnościowych przebywania SSP w wyróżnionych do rozważań stanach. Przedstawiona w artykule metodyka badania SSP ze względu na spełnienie określonych wymagań eksploatacyjnych może być użyta podczas opracowywania scenariusza pożarowego oraz projektowania systemów z uwzględnieniem różnych dostępnych rozwiązań technicznych.

***Słowa kluczowe**: eksploatacja, niezawodność, systemy sygnalizacji pożaru.*

1. Introduction

Transport fire alarm systems function in different, often extreme, operational condition. Type A, B fire alarm circuits, detection loops, alarm control units are located inside buildings (e.g. railway stations, signal boxes, transformer stations, switchgear), as well as outside, within an open access environment (e.g. railway stations, walkways, storage sheds, etc.) [10,11,12]. Long-term studies of the FAS operation process support the thesis that the proper functioning of these safety platforms is a reliability function of the components – sensors, modules, control units, etc. The maintenance and servicing process, spare parts availability and conducting periodic inspections also determine an appropriate reliability level [4, 6, 16, 17, 19]. The analysis of the operational phenomena, occurring within a FAS should take into account two important issues: the reliability approach already at the stage of developing the design fire scenario, as well as the efficient operational management of these complex technical objects – i.e., for example, service availability, conducting preventive inspections and the parameters of their environments [7, 12, 16]. For this purpose, the authors of the article selected two representative FAS, which are most commonly used within vast transport areas and attempted to imitate

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

the occurring operational phenomena in the form of developed research models [10, 11].

Fire alarm systems are ones of the most important electronic safety systems (frequently installed at facilities due to the applicable statutory and legal requirements), which are operated within vast transport areas. The proper functioning of safety platforms involves the implementation of a previously assumed transport process with an acceptable risk level of adverse external and internal impacts (e.g. fire, burglary, assault, terrorist attack, etc.). [6, 16, 19]. The unreliability of individual electronic devices and systems, as well as the mistakes of operators supervising the operational process in real-time can lead to the occurrence of safety hazard or unreliability states [5, 8, 18, 21]. The safety and risk theory answers, e.g. questions regarding the effects of failures, damage and operator mistakes. This is the cause for unacceptable states within these systems, e.g., unreliability or safety hazard. An important issue, which should be clarified by safety platform operators is the determination of a set of acceptable and unacceptable FAS states in terms of the safety of a given transport facility. [6, 16, 19, 23, 28, 29]. The scope of interest of the theory of safety and risk includes the results of damage and errors, which lead to the states of safety unreliability and hazard of the systems. The issue of correct clarification, which of the FAS states can be deemed acceptable or unacceptable from the perspective of safety or the initial fire scenario developed for a transport facility is very significant in this case [10, 11, 12, 16, 19].

A set of unacceptable states of an FAS can be reversible in the event of such a system having elements or devices, which initiate or interrupt a damage or failure process (including faulty operator actions) [6, 16, 19]. The counter-actions must be executed during the available time, where there is a possibility to remedy a dangerous situation [16]. Such actions are possible when safety platforms have a "reservoir" of permissible counter-failure operations. In such a case, it is impossible to move from permissible (e.g. surveillance) to prohibited (e.g. failure of a module or control unit – safety hazard state) states [6, 10, 11, 16, 19].

The service life of safety platforms in the case of adverse impacts can be increased by executing the available actions, e.g. by using redundancy or technical solutions improving the reliability of the devices themselves [10, 12, 16, 19]. A sensor(s) that uses numerous detectors reacting to the phenomenon of fire. Using redundancy means tolerating certain damage, as well as system expansion. The second case is preventing catastrophic damage, e.g. sensors within the system.

The quality of information [13, 14, 15, 16, 20] received by the systems from detectors [10, 11], installed over a vast transport area with a distorted electromagnetic environment (high levels of interfering signals) is also important. Certain research papers propose the application of fuzzy logic [22] or artificial neural networks [6, 13, 14, 15, 16, 23, 24], which are already used in detectors for developing alarm signals [10, 11]. The functionality of electronic transport systems is also significantly impacted by environmental conditions, temperature, humidity, vibrations and oscillations [3], as well as electromagnetic interference [1, 7, 18, 20, 21] but they are not included in this article. The article presents an operational and reliability analysis of a FAS, operated over a vast transport area. The analysis of the obtained results regarding the operational process, i.e., the measurement of times of restoration and damage occurrence enabled developing a FAS research model, and then, conducting an operational and reliability analysis taking into account the determined restoration and damage times [2, 8, 9, 16, 17, 19, 26, 28].

2. Representative transport fire alarm systems.

In the age of rapid technological progress and a constant development of the infrastructure, transport facilities are exposed to numerous hazards [6, 16, 19, 25, 26, 27]. The hazards not directly associated with fire, such as the threat of terrorism, can be its source [6, 16, 19]. This is why, a correct protection of transport facilities using active and passive fire safety equipment is an extremely important issue – Fig. 1.

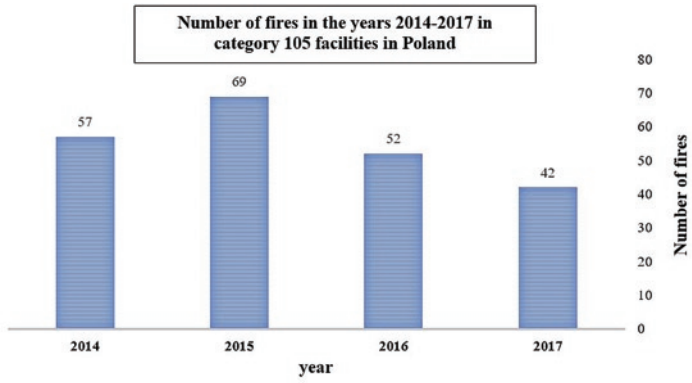


Fig. 1. Statistics regarding the number of fires in passenger service facilities, railway and bus stations, river and sea ports, and airports in particular, 2014-2017

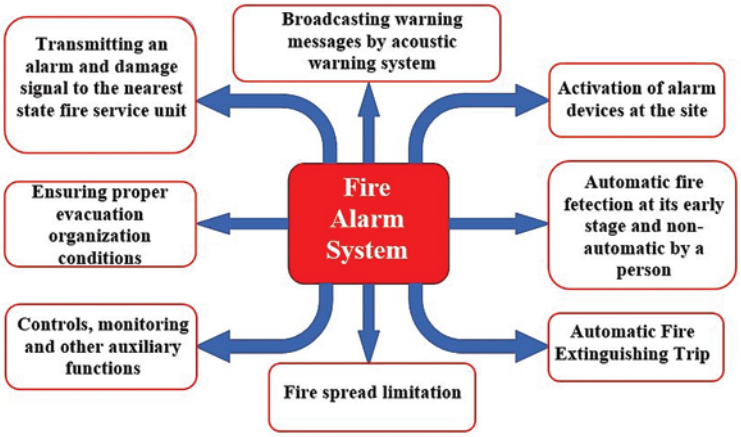


Fig. 2. Basic tasks executed by a fire alarm system.

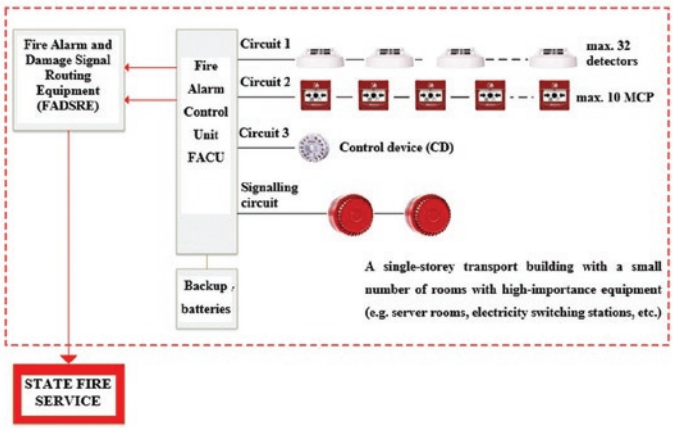


Fig. 3. Focused FAS with open detection circuits connected to an SFS fire signal and damage signal monitoring system (FACU – fire alarm control unit)

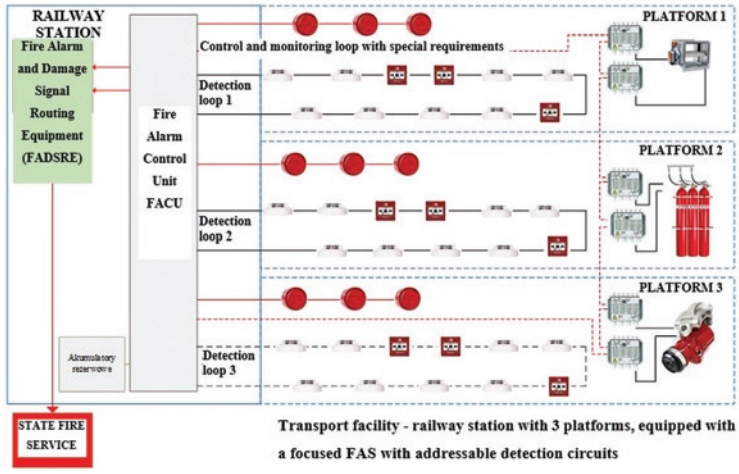


Fig. 4. Diagram of a focused fire alarm system with addressable detection circuits at a railway station with three platforms

According to the Regulation of the Minister of Interior and Administration (MSWiA) Dz. U. No. 109, item 719, by technical fire safety measures one should understand devices, equipment, systems and construction solutions designed to prevent the formation and spread of fires. The regulation understands fire safety equipment to be devices (fixed or semi-fixed, activated manually or automatically) aimed at preventing the formation of, detecting and fighting a fire or limiting its results. In particular, these include fixed and semi-fixed extinguishing and protecting devices, inerting devices, devices being part of an acoustic warning system (AWS) and FAS. FAS is a system, which includes signalling-alarm devices intended for automatic detection and transmittance of fire information, as well as receivers of fire alarms and receivers for damage signals – Figure 2 [6, 10, 11, 16, 19].

Several FAS types depending on the design, configuration and type of used linear elements are distinguished – Figure 3. The application of a given FAS type depends on the legal requirements for such systems, a fire scenario, which must be implemented, legal requirements for a given facility subject to protection, the adopted protection scope and the functional and utility requirements, which must be satisfied by the system. The fire origin (fire source) location indication accuracy depends on the used FAS. On the other hand, the requirement regarding the fire location accuracy is a criterion for the selection of a fire alarm system [10, 12].

In the case of a conventional (non-addressable) FAS, indicating the fire detection location, is limited to the detection circuit, whereas in the case of an addressable system, the control unit indicates the fire source location with an accuracy down to a fire detector (depending on the configuration, down to a fire zone) [11, 12]. The type of an FAS installed at a transport facility impacts its division into the so-called detection zones. The control and monitoring loop must be executed in accordance with the special requirements, and in a manner so as to maintain the power supply or signal transmission continuity for the time period required for device commissioning and operation, pursuant to §187 art. 2 of the Regulation by the Minister of Infrastructure of 12 April 2002 (Dz. U. No. 75, item 690, as amended) [10, 11, 12, 16].

Due to the small extent of the transport facility, short distance of the loop cabling routes and a low number of controls and monitoring devices, often a single control and monitoring loop is used, which can handle e.g. all platforms – Figure 4.

3. The analysis of selected fire alarm system reliability and operational process

A focused fire alarm system based on a conventional fire alarm control unit with a single detection open circuit equipped with a maximum of 32 fire detectors and a signalling circuit with two sounders is shown in Figure 5. Whereas Figure 6 shows the relationships occurring within a focused system with a fire alarm control unit with connected open detection circuit with optical smoke detectors and a signalling circuit with sounders.

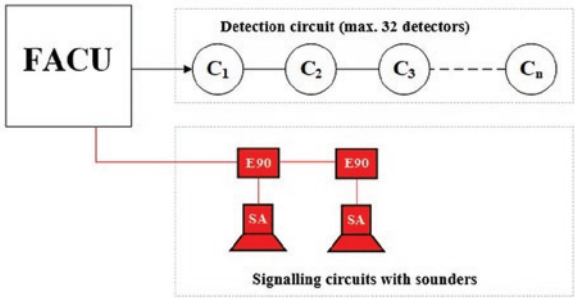


Fig. 5. Focused FAS with an open detection circuit and a signalling circuit with sounders

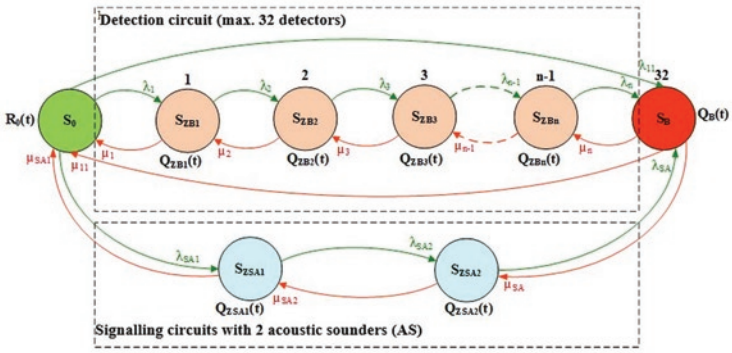


Fig. 6. Relationships occurring within a focused system with a fire alarm control, unit with connected open detection circuit with optical smoke detectors and a signalling circuit with sounders

Relationships occurring within a system – fig. 6 can be described with the following relations [10,19] (1):

$$
\begin{aligned}
R_0'(t) &= -\lambda_{11}\cdot R_0(t) - \lambda_1\cdot R_0(t) - \lambda_{SA1}\cdot R_0(t) + \mu_{11}\cdot Q_B(t) + \mu_1\cdot Q_{ZB1}(t) + \mu_{SA1}\cdot Q_{ZSA1}(t)\\
Q'_{ZB1}(t) &= -\lambda_2\cdot Q_{ZB1}(t) - \mu_1\cdot Q_{ZB1}(t) + \lambda_1\cdot R_0(t) + \mu_2\cdot Q_{ZB2}(t)\\
Q'_{ZB2}(t) &= -\lambda_3\cdot Q_{ZB2}(t) - \mu_2\cdot Q_{ZB2}(t) + \lambda_2\cdot Q_{ZB1}(t) + \mu_3\cdot Q_{ZB3}(t)\\
Q'_{ZB3}(t) &= -\lambda_{n-1}\cdot Q_{ZB3}(t) - \mu_3\cdot Q_{ZB3}(t) + \lambda_3\cdot Q_{ZB2}(t) + \mu_{n-1}\cdot Q_{ZBn}(t)\\
&\text{----------}\\
Q'_{ZBn}(t) &= -\lambda_n\cdot Q_{ZBn}(t) - \mu_{n-1}\cdot Q_{ZBn}(t) + \lambda_{n-1}\cdot Q_{ZB3}(t) + \mu_n\cdot Q_B(t)\\
Q'_{ZSA1}(t) &= -\lambda_{SA2}\cdot Q_{ZSA1}(t) - \mu_{SA1}\cdot Q_{ZSA1}(t) + \lambda_{SA1}\cdot R_0(t) + \mu_{SA2}\cdot Q_{ZSA2}(t)\\
Q'_{ZSA2}(t) &= -\lambda_{SA}\cdot Q_{ZSA2}(t) - \mu_{SA2}\cdot Q_{ZSA2}(t) + \lambda_{SA2}\cdot Q_{ZSA1}(t) + \mu_{SA}\cdot Q_B(t)\\
Q'_B(t) &= -\mu_{11}\cdot Q_B(t) - \mu_n\cdot Q_B(t) - \mu_{SA}\cdot Q_B(t) + \lambda_{11}\cdot R_0(t) + \lambda_n\cdot Q_{ZBn}(t) + \lambda_{SA}\cdot Q_{ZSA2}(t)
\end{aligned}
\tag{1}
$$

Adopting the baseline conditions [10, 19] (2):

$$R_0(t) = 1$$

$$Q_{ZB1}(0) = Q_{ZB2}(0) = Q_{ZB3}(0) = ... = Q_{ZBn}(0) = Q_B(0) = \\ = Q_{ZSA1}(0) = Q_{ZSA2}(0) = 0 \quad (2)$$

where [10,19]:

- $R_0(t)$ – probability function of the system staying in the state of full fitness S_{PZ};
- $Q_{ZB1}(t)$, $Q_{ZBn}(t)$, $Q_{ZSA1}(t)$, $Q_{ZSA2}(t)$ – probability function of the system staying in individual safety hazard states;
- $Q_B(t)$ – probability function of the system staying in the state of safety unreliability S_B;
- λ_{11} – intensity of transition from the state of full fitness S_{PZ} to the state of safety unreliability S_B;
- μ_{11} – intensity of transition from the state of safety unreliability S_B to the state of full fitness S_{PZ};
- λ_1, λ_1,… – intensity of transitions from the state of full fitness S_{PZ} or the state of safety hazard $S_{ZB1,2,...}$ to the state of safety unreliability $Q_B(t)$ or the state safety hazard or the state of safety reliability S_{ZB} – according to the designation in Figure 6;
- μ_1, μ_2, … – intensity of transitions from the state of safety hazard S_{ZB} to the state of full fitness S_{PZ}, from the state of safety unreliability to the state of safety hazard Q_{ZBn}, Q_{ZB}, Q_{ZB2}, … – according to the designations as in Figure 6.

Figure 7 shows a focused FAS, based on an addressable FACU, with connected open detector and manual call point circuits. All elements are equipped with short-circuit isolators. The system consists of loop circuits, some of which have programmed detectors in coincidence systems, a control loop with a module controlling fire safety devices, as well as technical and safety systems within the signal box room. A signalling circuit with sounders is also hooked in to the control unit [6, 10, 11, 12].

Figure 7 shows the relationships occurring within a focused system with an addressable fire alarm control unit with looped open circuits and a signalling circuit. The system presented in Fig. 7 can be described by the following Chapman–Kolmogorov equations (3). The relationships occurring within a focused system are shown in Fig. 8.

Owing to the various structures of FASs operated within a vast transport area, the relationships between individual devices in the system can differ, which is shown in Figures 6 and 8. The

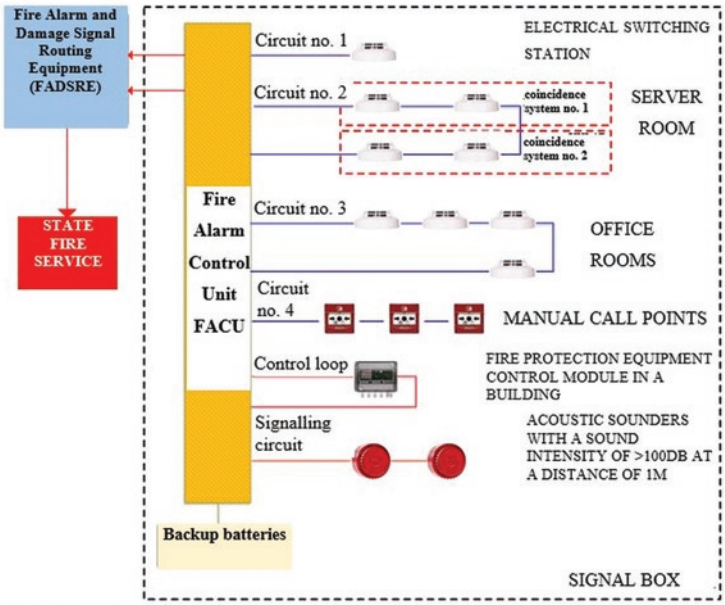


Fig. 7. Focused FAS with addressable fire alarm control units with open, looped circuits and a signalling circuit

Fig. 7. Focused FAS with addressable fire alarm control units with open, looped circuits and a signalling circuit

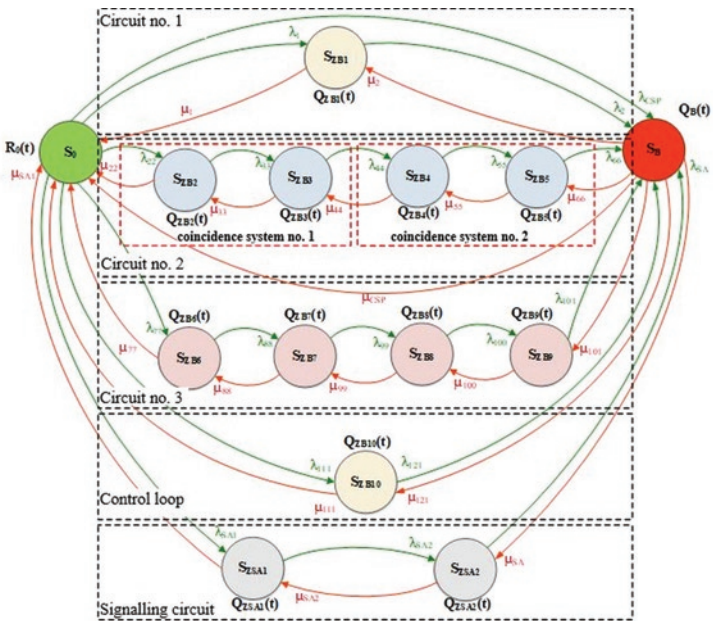


Fig. 8. Relationships occurring within a focused system with addressable fire alarm control unit with open, looped circuits and a signalling circuit

Table 1. Studying the operational process of FASs used in transport facilities

No.	Damage type	Failure time	Failure removal time	Repair duration	Repair type
1	Circuit no. 3 interference	3/1/2018 14:32	03/01/2018 18:10	3h 38 min.	Improving the tie-in of the circuit to control unit terminals and control unit reset
2	Failure of detector 3/57	1/2/2018 18:10	2/2/2018 23:30	5h 20 min.	Replacing detector with a new one
n-2	Circuit no. 1 interference	1/12/2018 04:15	1/12/2018 09:00	4h 45 min.	Central unit reset
n-1	Failure of detector 3/11	15/12/2018 11:15	15/12/2018 14:20	4h 5 min.	Replacing detector socket
n	CSO2 communication error	27/12/2018 15:05	28/12/2018 09:05	18h 5 min.	CSO2 and fire alarm control unit

system shown in Figure 7 has a more complex reliability structure due to the presence of more fire-protected objects – server rooms, office areas and an electrical switching station. For this reason, separate detection circuits 1, 2 and 3, a control loop and a signalling circuit can be distinguished. In addition, the detection circuit no. 2 utilizes alarms in the coincidence system. In the case of such a designed FAS, one should distinguish more operating states, which is why the system of equations (3) describing the behaviour of the system within the operational process becomes complex.

$$
\begin{gathered}
R_0'(t) = -\lambda_{CSP} \cdot R_0(t) - \lambda_1 \cdot R_0(t) - \lambda_{22} \cdot R_0(t) - \lambda_{77} \cdot R_0(t) - \lambda_{111} \cdot R_0(t) - \lambda_{SA1} \cdot R_0(t) + \\
+ \mu_{CSP} \cdot Q_B(t) + \mu_1 \cdot Q_{ZB1}(t) + \mu_{22} \cdot Q_{ZB2}(t) + \mu_{77} \cdot Q_{ZB6}(t) + \mu_{111} \cdot Q_{ZB10}(t) + \mu_{SA1} \cdot Q_{ZSA1}(t) \\
Q'_{ZB1}(t) = -\mu_1 \cdot Q_{ZB1}(t) - \lambda_2 \cdot Q_{ZB1}(t) + \mu_2 \cdot Q_B(t) + \lambda_1 \cdot R_0(t) \\
Q'_{ZB2}(t) = -\mu_{22} \cdot Q_{ZB2}(t) - \lambda_{33} \cdot Q_{ZB2}(t) + \mu_{33} \cdot Q_{ZB3}(t) + \lambda_{22} \cdot R_0(t) \\
Q'_{ZB3}(t) = -\mu_{33} \cdot Q_{ZB3}(t) - \lambda_{44} \cdot Q_{ZB3}(t) + \mu_{44} \cdot Q_{ZB4}(t) + \lambda_{33} \cdot Q_{ZB2}(t) \\
Q'_{ZB4}(t) = -\mu_{44} \cdot Q_{ZB4}(t) - \lambda_{55} \cdot Q_{ZB4}(t) + \mu_{55} \cdot Q_{ZB5}(t) + \lambda_{44} \cdot Q_{ZB3}(t) \\
Q'_{ZB5}(t) = -\mu_{55} \cdot Q_{ZB5}(t) - \lambda_{66} \cdot Q_{ZB5}(t) + \mu_{66} \cdot Q_B(t) + \lambda_{55} \cdot Q_{ZB4}(t) \\
Q'_{ZB6}(t) = -\mu_{77} \cdot Q_{ZB6}(t) - \lambda_{88} \cdot Q_{ZB6}(t) + \mu_{88} \cdot Q_{ZB7}(t) + \lambda_{77} \cdot R_0(t) \\
Q'_{ZB7}(t) = -\mu_{88} \cdot Q_{ZB7}(t) - \lambda_{99} \cdot Q_{ZB7}(t) + \mu_{99} \cdot Q_{ZB8}(t) + \lambda_{88} \cdot Q_{ZB6}(t) \\
Q'_{ZB8}(t) = -\mu_{99} \cdot Q_{ZB8}(t) - \lambda_{100} \cdot Q_{ZB8}(t) + \mu_{100} \cdot Q_{ZB9}(t) + \lambda_{99} \cdot Q_{ZB7}(t) \\
Q'_{ZB9}(t) = -\mu_{100} \cdot Q_{ZB9}(t) - \lambda_{101} \cdot Q_{ZB9}(t) + \mu_{101} \cdot Q_B(t) + \lambda_{100} \cdot Q_{ZB8}(t) \\
Q'_{ZB10}(t) = -\mu_{111} \cdot Q_{ZB10}(t) - \lambda_{121} \cdot Q_{ZB10}(t) + \mu_{121} \cdot Q_B(t) + \lambda_{111} \cdot R_0(t) \\
Q'_{ZSA1}(t) = -\mu_{SA1} \cdot Q_{ZSA1}(t) - \lambda_{SA2} \cdot Q_{ZSA1}(t) + \mu_{SA2} \cdot Q_{ZSA2}(t) + \lambda_{SA1} \cdot R_0(t) \\
Q'_{ZSA2}(t) = -\mu_{SA2} \cdot Q_{ZSA2}(t) - \lambda_{SA} \cdot Q_{ZSA2}(t) + \mu_{SA} \cdot Q_B(t) + \lambda_{SA2} \cdot Q_{ZSA1}(t) \\
Q'_B(t) = -\mu_{CSP} \cdot Q_B(t) - \mu_2 \cdot Q_B(t) - \mu_{66} \cdot Q_B(t) - \mu_{101} \cdot Q_B(t) - \mu_{121} \cdot Q_B(t) - \mu_{SA} \cdot Q_B(t) + \\
+ \lambda_{CSP} \cdot R_0(t) + \lambda_2 \cdot Q_{ZB1}(t) + \lambda_{66} \cdot Q_{ZB5}(t) + \lambda_{101} \cdot Q_{ZB9}(t) + \lambda_{121} \cdot Q_{ZB10}(t) + \lambda_{SA} \cdot Q_{ZSA2}(t)
\end{gathered}
\tag{3}
$$

Adopting the baseline conditions (4):

$$
\begin{gathered}
R_0(t) = 1 \\
Q_{ZB1}(0) = Q_{ZB2}(0) = Q_{ZB3}(0) = Q_{ZB4}(0) = Q_{ZB5}(0) = Q_{ZB6}(0) = Q_{ZB7}(0) = \\
= Q_{ZB8}(0) = Q_{ZB9}(0) = Q_{ZB10}(0) = Q_{ZSA1}(0) = Q_{ZSA2}(0) = Q_B(0) = 0
\end{gathered}
$$

4. Operational statistics (repairs, damage) regarding representative FAS

The analysis in the scope of FAS operational process was conducted for n=20 various systems. The structure of the studied FAS corresponded to the fire safety systems used at transport facilities. The FAS operational studies covered: restoration, damage occurrence time and repair time. The studies were conducted for the following FAS types operated within a transport area:

a) FAS with an addressable fire alarm control unit and one detection loop (n=15 units);
b) FAS with an addressable fire alarm control unit and two detection loops (n=3 units);
c) FAS with an addressable FACU, three detection loops, one control-detection loop for monitoring fixed extinguishing devices and generating their tripping signal (n=2 units).

All of the aforementioned FAS were operated in similar environmental conditions (temperature, humidity, pressure, etc.) in transport buildings. Owing to the importance of FAS in ensuring the transport process safety, the service team dealing with repairs and restorations was available within 2 hours from the damage being reported by persons monitoring the operation (for n=15 FAS). Other systems (n=5) had the damage report response time extended to 4 hours due to the supervision over transport facilities – buildings, which do not pose a direct threat for the passenger transport (e.g. warehouses, sheds, etc.). Tables 1-3 show examples of the FAS operational process study results.

Table 1 shows representative types of damage for selected FAS. The data were compiled based on a set of damage for n = 20 FAS, operated over a vast transport area. A maximum repair time was adopted for a given type of FAS damage (n=20 units). The repair time does not include the service personnel travel time (in the case of such FAS, such personnel should be on site).

5. FAS operational process modelling in the RELIASOFT BLOCKSIM software

Calculations involving the probability for a system to be in the states of safety hazard, safety unreliability and full fitness for the FAS operational process model were conducted using a commercial, specialized computational BlockSim software by ReliaSoft. The computations were conducted for a focused FAS model – open circuits, no notifications. Tables 4 and 5 show the calculated parameters, e.g. initial and average probability, availability coefficient for time t for individual states or time the FAS spends in a given state.

For the example operating time t = 4 201 h, the values of the availability coefficient $K_g(t)$ for individual states S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2} of a fire alarm system are shown in table 6, and the percentage share of FAS in individual states in Figure 12. Figure 13 shows the growth rate of state availability coefficient for a selected FAS operating interval.

Table 2. Repair time with the maximum time T_{max}, annualized

No.	Repair of a given damage type	Failure time	Failure removal time	Maximum repair time [T_{max}]
Detection loop 1 damage				
	Circuit no. 3 interference	3/1/2018 14:32	3/1/2018 18:10	3h 38 min.
	Circuit no. 2 interference	11/3/2018 15:00	11/3/2018 16:30	1h 30 min.
	Ground fault of loop no. 1	2/5/2018 13:30	2/5/2018 19:00	5h 30 min.
	Circuit no. 1 interference	1/12/2018 04:15	1/12/2018 09:00	4h 45 min.
	Communication error loop 1	30/11/2018 10:30	30/11/2018 14:30	4h
Manual call point damage				
	Failure of MCP 1/10	15/6/2018 09:20	15/6/2018 13:20	4h
	Failure of MCP 1/10	16/6/2018 14:00	16/6/2018 19:05	5h 5 min.
FAS power supply failure				
	230V power failure	27/2/2018 11:30	27/2/2018 11:45	15 min.
	FACU battery failure	16/4/2018 19:00	17/4/2018 08:10	13h 10 min.

Table 3. Damage intensity along with marked times of fire alarm system unfitness, annualized (example).

No.	Repair of a given damage type	Failure time	Failure removal time	Repair time [T_{max}]
Detection loop damage				
1.	Circuit no. 3 interference	3/1/2018 14:32	3/1/2018 18:10	3h 38 min.
2.	Ground fault of loop no. 1	2/5/2018 13:30	2/5/2018 19:00	5h 30 min.
3.	Circuit no. 1 interference	1/12/2018 04:15	1/12/2018 09:00	4h 45 min.
4.	Communication error loop 1	30/11/2018 10:30	30/11/2018 14:30	4h
Total FAS unfitness time, annualized:				19h 23 min.
FAS power supply failure				
1.	230V power failure	27/2/2018 11:30	27/2/2018 11:45	15 min.
2.	FACU battery failure	16/4/2018 19:00	17/4/2018 08:10	13h 10 min.
Total FAS unfitness time, annualized:				13h 25 min.

Table 4. Fire alarm system parameters for time t = 8 760 [h]

Name of state	Initial probability	Average probability	Availability for time t [8760 h]	Reliability for time t	Time in a given state
S_0	1	0,999993444	0,999993439	0,991489928	8759,94257
S_B	0	2,245 E-07	2,24528 E-07	0,001526641	0,001966621
S_{ZBI}	0	3,75408 E-06	3,75731 E-06	0,003920964	0,032885763
S_{ZBI2}	0	8,25865 E-07	8,26355 E-07	0,001033117	0,00723458
S_{ZBI3}	0	7,10979 E-07	7,11386 E-07	0,000996234	0,006228174
S_{ZBP1}	0	8,16121 E-07	8,16726 E-07	0,001033117	0,007149221
S_{ZBP2}	0	2,24374 E-07	2,24529 E-07	0	0,001965516

Table 5. Intensity matrix for individual FAS states for t = 8 760 [h]

From → to	S_0	S_B	S_{ZBI}	S_{ZBI2}	S_{ZBI3}	S_{ZBP1}	S_{ZBP2}
S_0	-	1,7502 E-07	4,49514 E-07	1,1844 E-07	1,14212 E-07	1,1844 E-07	0
S_B	0,0759	-	0,1818	0,1968	0,125	0	0,2
S_{ZBI}	0,1305	2,52906 E-07	-	0	0	0	0
S_{ZBI2}	0,1968	5,70919 E-08	0	-	0	0	0
S_{ZBI3}	0,2	1,4161 E-08	0	0	-	0	0
S_{ZBP1}	0,2	0	0	0	0	-	1,18 E-07
S_{ZBP2}	0	1,4161 E-08	0	0	0	0.2	-

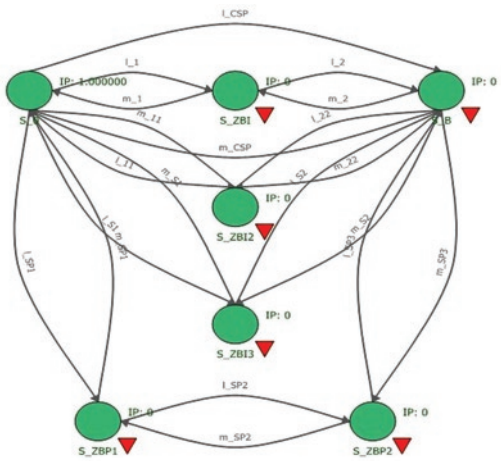


Fig. 9. Migration of possible focused FAS with open circuits, without notifying the SFS (where: SFS – State Fire Service).

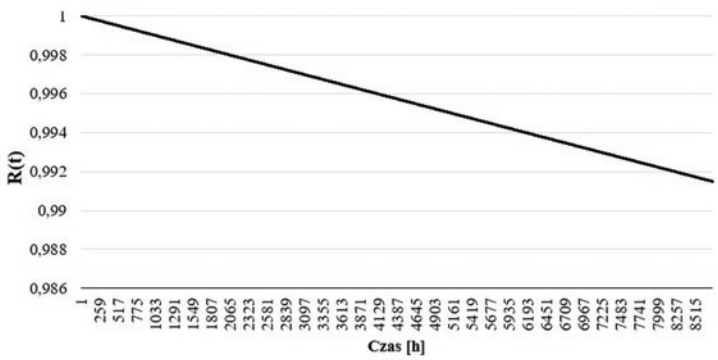


Fig. 10. Reliability R(t) of a FAS with open circuits, without SFS notifications

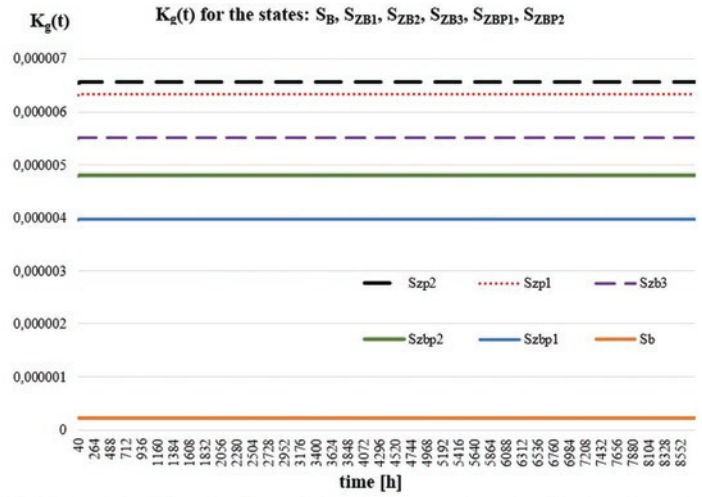


Fig. 11. Zonal (partial) availability coefficient for the states of S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2} SSP (safety unreliability and hazard); the graph does not show state S_0 (for $t = 0$ $S_0(t) = 1$)

The growth rate of a zonal availability coefficient for an individual state can be expressed using the formula (5):

$$S_{ZB1} = \frac{\Delta Kg_{SZB1}}{\Delta t}\left[\frac{1}{h}\right]$$
$$S_{ZB1} = \frac{(2{,}75466E\text{-}6) - (2{,}27438E\text{-}6)}{10{,}25 - 7{,}25} = 1{,}08915E\text{-}10\left[\frac{1}{h}\right] \quad (5)$$

The growth rate values for the availability coefficient K_g for other fire alarm system states are shown in table 7 and Figure 14. Fig. 15 shows the probability of a FAS staying in individual states.

Growth rate R(t) over time Δt for a particular state is described using the formula (6):

$$S_{ZB1} = \frac{\Delta R(t)_{SZB1}}{\Delta t}\left[\frac{1}{h}\right]$$
$$S_{ZB1} = \frac{(2{,}75466E - 6) - (2{,}27438E - 6)}{10{,}25 - 7{,}25} = 1{,}08915E - 10\left[\frac{1}{h}\right] \quad (6)$$

Tab. 6. Values of $K_g(t)$ coefficient for individual FAS over time

Time [h]	Fire alarm system state					
	S_B	S_{ZB1}	S_{ZB2}	S_{ZB3}	S_{ZBP1}	S_{ZBP2}
	Value of coefficient $K_g(t)$					
4 201	$2245282 \bullet 10^{-6}$	$3757312 \bullet 10^{-5}$	$82635535 \bullet 10^{-6}$	$71138635 \bullet 10^{-6}$	$81672633 \bullet 10^{-6}$	$2245287 \bullet 10^{-6}$

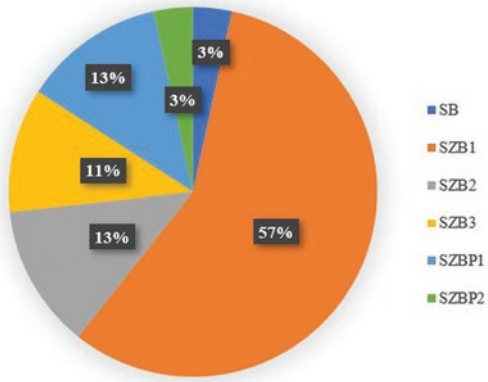


Fig. 12. Percentage share of FAS staying in a given state, according to table 6

Tab. 7. Growth rate of availability coefficients K_g for individual FAS states.

No.	FAS state	Growth rate S of the availability coefficient K_g [1/h]
1.	S_B	1,37487E-10
2.	S_{ZB1}	2,08915E-10
3.	S_{ZB2}	4,99493E-11
4.	S_{ZB3}	2,2907E-11
5.	S_{ZBP1}	1,05053E-11
6.	S_{ZBP2}	4,81767E-12

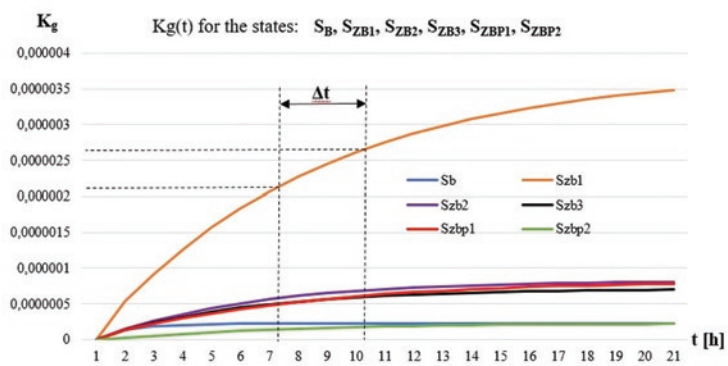


Fig. 13. Growth rate for the availability coefficients of states S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2} of a fire alarm system (safety unreliability and hazard); the graph does not show the state S_0 (for $t = 0$, $K_g(t) = 1$), adopted time $t = 21$ h in order to reflect the parameter change rates at the initial stage of transition state changes

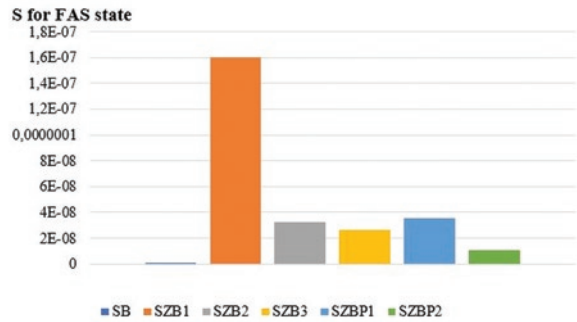


Fig. 14. Growth rate for the availability coefficients of states S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2}

6. Conclusions

Fire alarm systems operated over vast transport areas have various connection structures, which are a function of the executed tasks – fire monitoring of buildings [6, 10, 11, 12, 16].

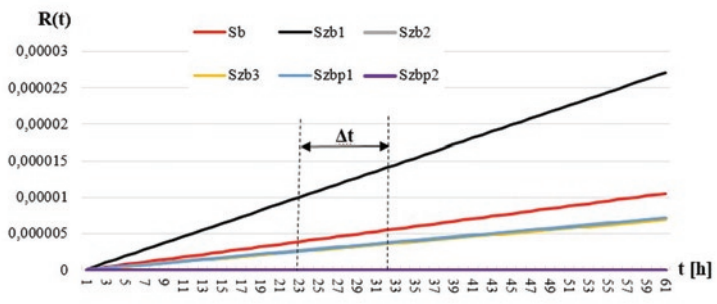


Fig. 15. Probability of an FAS staying in a state R(t) for states S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2}; the graph does not show R(t) for the state S_0 (for t = 0, R(t) = 1), adopted time t = 61 h in order to depict the parameter change rate at the initial change phase

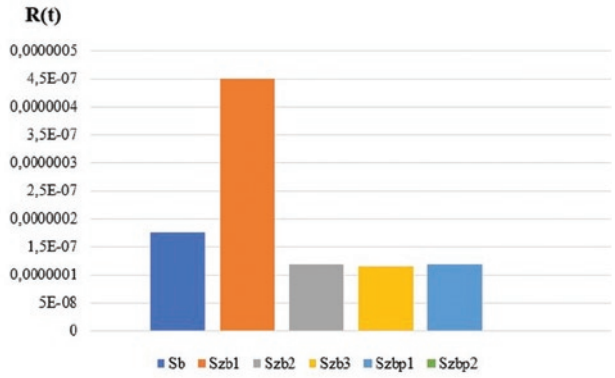


Fig. 16. Growth rate R(t) values over time for selected FAS states

Complex FAS have a dozen or so detection loops, as well as signaller, desmoking control, gas suppression, etc. lines. Owing to the extent of the executed tasks and fire controls, the reliability and operational structure of such systems is mixed. Available technical measures are applied in order to increase FAS reliability. The article presents a model and operational and reliability analysis of a selected FAS, which is operated within a transport area. Seven operating states were distinguished for the system. The average value for the probability of a system staying in the state of fitness was S0 = 0,999993444, whereas the time spent in this state was 8759, 94257 [h] (the simulation time was t = 1 year of FAS operation). When considering the so-called Kgs(t) zonal (partial) availability coefficients for FAS states of SB, SZB1, SZB2, SZB3, SZBP1, SZBP2 it can be observed, that the SZBP2 state is dominant at the initial operational stage. Therefore, when designing a FAS, particular attention to the transition between the states of fitness S0 and the state of safety hazard SZBP2 should be paid. In order to depict the parameter change rates at the initial stage of transition state changes, the growth rate for the availability coefficients for FAS states of SB, SZB1, SZB2, SZB3, SZBP1, SZBP2 was determined. At the initial FAS operational stage, the highest value was obtained for the state SZB1 = 2,08915E-10 [1/h]. All zonal (so-called partial) availability coefficients stabilize their values throughout the further periods of the operational process – Fig. 1. The R(t) probability of a FAS staying in a state is very low for individual states SB, SZB1, SZB2, SZB3, SZBP1, SZBP2 during the initial operational period – Fig. 15. The highest growth rate of the R(t) value during the initial operational process was exhibited by FAS state SZB1.

References

1. Białek K, Paś J. Exploitation of selected railway equipment - conducted disturbance emission examination, Diagnostyka 2018; 19(3): 29-35, https://doi.org/10.29354/diag/92003.
2. Branson D. Stirling numbers and Bell numbers, their role in combinatorics and probability, Math. Scientist 2000; 25: 1-31.
3. Burdzik R, Konieczny Ł, Figlus T. Concept of on-board comfort vibration monitoring system for vehicles, In the monograph Activities of Transport Telematics, editors: Mikulski J, Springer 2013: 418-425, https://doi.org/10.1007/978-3-642-41647-7_51.
4. Duer S, Zajkowski K, Duer R, Paś J. Designing of an effective structure of system for the maintenance of a technical object with the using information from an artificial neural network, Neural Computing & Applications 2012; 23(3): 913-925, https://doi.org/10.1007/s00521-012-1016-0.
5. Duer S, Scaticailov S, Paś J, Duer R, Bernatowicz D. Taking decisions in the diagnostic intelligent systems on the basis information from an artificial neural network, 22nd International Conference on Innovative Manufacturing Engineering and Energy, MATEC Web of Conferences 178, 2018; 178: 1-6, https://doi.org/10.1051/matecconf/201817807003.
6. Dyduch J, Paś J, Rosiński A. The basic of the exploitation of transport electronic systems, Radom: Publishing House of Radom University of Technology, 2011.
7. Dziula P, Paś J. Low Frequency Electromagnetic Interferences Impact on Transport Security Systems Used in Wide Transport Areas, TransNav the International Journal on Marine Navigation and Safety of Sea Transportation 2018, 12(2): 251-258, https://doi.org/10.12716/1001.12.02.04.
8. Garmabaki A H S, Ahmadi A, Mahmood Y A, Barabadi A. Reliability modelling of multiple repairable units, Quality and Reliability Engineering International 2016; 32(7): 2329-2343, https://doi.org/10.1002/qre.1938.
9. Jachimowski R, Żak J, Pyza D. Routes planning problem with heterogeneous suppliers demand, 21st International Conference on Systems Engineering, Las Vegas, USA 2011: 434-437, https://doi.org/10.1109/ICSEng.2011.85.
10. Klimczak T, Paś J. Analysis of reliability structures for fire signaling systems in the field of fire safety and hardware requirements, Journal of KONBIN 2018; (64): 191-214, https://doi.org/10.2478/jok-2018-0030.
11. Klimczak T, Paś J. Analysis of solution of a fire signaling system for a choice railway building, Biuletyn WAT 2018; (67)4: 195-205, https://doi.org/10.5604/01.3001.0012.8515.
12. Klimczak T, Paś J. Electromagnetic environment on extensive logistic areas and the proces of using electronic safety system, Politechnika Warszawska, Prace Naukowe Transport 2018; (121): 135-146.
13. Krzykowski M, Paś J, Rosiński A. Assessment of the level of reliability of power supplies of the objects of critical infrastructure, IOP Conf. Series: Earth and Environmental Science 2019: 1-9, https://doi.org/10.1088/1755-1315/214/1/012018.
14. Lewiński A, Perzyński T, Toruń A. The analysis of open transmission standards in railway control and management, Communications in Computer and Information Science, Berlin Heidelberg Springer-Verlag, 2012; 329: 10-17, https://doi.org/10.1007/978-3-642-34050-5_2.

15. Łubkowski P, Laskowski D. Selected issues of reliable identification of object in transport systems using video monitoring services, Communication in Computer and Information Science, Berlin Heidelberg Springer, 2015, Vol. 471: 59-68.https://doi.org/10.1007/978-3-662-45317-9_7
16. Paś J. Operation of electronic transportation systems. Radom: Publishing House University of Technology and Humanities, 2015.
17. Paś J, Duer S. Determination of the impact indicators of electromagnetic interferences on computer information systems. Neural Computing & Applications 2012; 23(7): 2143-2157, https://doi.org/10.1007/s00521-012-1165-1.
18. Paś J, Rosiński A. Selected issues regarding the reliability-operational assessment of electronic transport systems with regard to electromagnetic interference. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19(3): 375-381, https://doi.org/10.17531/ein.2017.3.8.
19. Rosiński A. Modelling the maintenance process of transport telematics systems. Warsaw: Publishing House Warsaw University of Technology, 2015.
20. Siergiejczyk M, Paś J, Rosiński A. Train call recorder and electromagnetic interference, Diagnostyka, 2015, 16(1): 19-22.
21. Siergiejczyk M, Paś J, Rosiński A. Issue of reliability-exploitation evaluation of electronic transport systems used in the railway environment with consideration of electromagnetic interference, IET Intelligent Transport Systems 2016;10(9): 587-593, https://doi.org/10.1049/iet-its.2015.0183.
22. Skorupski J, Uchroński P. A fuzzy reasoning system for evaluating the efficiency of cabin luggage screening at airports, Transportation Research Part C - Emerging Technologies 2015; (54): 157-175, https://doi.org/10.1016/j.trc.2015.03.017.
23. Stawowy M. Model for information quality determination of teleinformation systems of transport, In: Proceedings of the European Safety and Reliability Conference ESREL 2014, CRC Press/Balkema 2015: 1909-1914, https://doi.org/10.1201/b17399-261.
24. Stawowy M, Kasprzyk Z. Identifying and simulation of status of an ICT system using rough sets, Tenth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX, given as the monographic publishing series - "Advances in intelligent systems and computing", Springer 2015; 365: 477-484, https://doi.org/10.1007/978-3-319-19216-1_45.
25. Warczek J, Młyńczak J, Celiński I. Simulation studies of a shock absorber model proposed under conditions of different kinematic input functions, Vibroengineering Procedia 6, 2015: 248-253.
26. Weintrit A, Dziula P, Siergiejczyk M, Rosiński A. Reliability and exploitation analysis of navigational system consisting of ECDIS and ECDIS back-up systems, The monograph Activities in Navigation - Marine Navigation And Safety Of Sea Transportation, editors: Weintrit A, London: CRC Press/Balkema 2015: 109-115, https://doi.org/10.1201/b18513-17.
27. Weintrit A. Technical infrastructure to support seamless information exchange in e-Navigation, In: Mikulski, J. (ed.), TST 2013, Springer Heidelberg, CCIS 2013; 395: 188-199, https://doi.org/10.1007/978-3-642-41647-7_24.
28. Yang L, Yan X. Design for Reliability of Solid State Lighting Products, In: Solid State Lighting Reliability, eds: van Driel W, Fan X., Solid State Lighting Technology and Application Series, Springer, New York 2013; 1: 497-556, https://doi.org/10.1007/978-1-4614-3067-4_19.
29. Zajkowski K, Rusica I, Palkova Z. The use of CPC theory for energy description of two nonlinear receivers, MATEC Web of Conferences 2018; 178: 1-6, https://doi.org/10.1051/matecconf/201817809008.

Jacek PAŚ
Faculty of Electronics
Military University of Technology
ul. gen. S. Kaliskiego 2, 00-908 Warsaw, Poland

Tomasz KLIMCZAK
The Main School of Fire Service
ul. Juliusza Słowackiego 52/54, 01-629 Warsaw, Poland

E-mails: jacek.pas@wat.edu.pl, tklimczak@sgsp.edu.pl

Article citation info:
KUCHARSKA B. Identification of surface stress in the exhaust system pipe made by hydroforming technology based on diffractometric measurements. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 562–566, http://dx.doi.org/10.17531/ein.2019.4.4.

Barbara KUCHARSKA

IDENTIFICATION OF SURFACE STRESS IN THE EXHAUST SYSTEM PIPE MADE BY HYDROFORMING TECHNOLOGY BASED ON DIFFRACTOMETRIC MEASUREMENTS

IDENTYFIKACJA NAPRĘŻEŃ POWIERZCHNIOWYCH W RURZE DO UKŁADU WYDECHOWEGO WYKONANEJ TECHNOLOGIĄ HYDROFORMOWANIA NA PODSTAWIE POMIARÓW DYFRAKTOMETRYCZNYCH*

In the work identificaton of surface stresses in the exhaust pipe made of Cr-Ni steel shaped with hydroforming technology. Stresses were determined by the non-destructive x-ray method $\sin^2\psi$. A complex state of tensile stresses with values in the range of 69÷240 MPa for circumferential stresses and 26÷290 MPa for longitudinal stresses was found on the surface of the pipe. The distribution of stresses on the circumference and length of the pipe was analyzed on the basis of coefficients of variation and wall thickness. A relationship was found between the value of surface stress and the wall thickness of the pipe. The highest stresses occurred in the areas of the pipe where the thickness of the wall was reduced the most. In the central part of the pipe, where the wall thickness reduction was the smallest, the stresses were also the smallest, but they were characterized by the highest dispersion of value. The distribution of surface stresses determined by diffractometric method was compared with the model of deformation of the pipe generated numerically.

***Keywords**: hydroforming, exhaust system, surface stresses, x-ray stress measurement.*

W pracy dokonano identyfikacji naprężeń powierzchniowych w rurze wydechowej ze stali Cr-Ni kształtowanej technologią hydroformowania. Naprężenia wyznaczono nieniszczącą rentgenowską metodą $\sin^2\psi$. Na powierzchni rury stwierdzono złożony stan naprężeń rozciągających o wartościach z zakresu 69÷240 MPa dla naprężeń obwodowych i 26÷290 MPa dla naprężeń wzdłużnych. Rozłożenie naprężeń na obwodzie i długości rury analizowano na podstawie współczynników zmienności i grubości ścianki. Stwierdzono zależność pomiędzy wartością naprężeń powierzchniowych a grubością ścianki rury. Największe naprężenia występowały w obszarach rury gdzie grubość ścianki była najsilniej zredukowana. W centralnej części rury gdzie redukcja grubości ścianki była najmniejsza naprężenia również były najmniejsze, ale cechowały się największym rozproszeniem wartości. Rozłożenie naprężeń powierzchniowych wyznaczonych metodą dyfraktometryczną porównano z modelem odkształceń w rurze wygenerowanym numerycznie.

***Słowa kluczowe**: hydroformowanie, układ wydechowy, naprężenia powierzchniowe, rentgenowski pomiar naprężeń.*

1. Introduction

For the production of exhaust systems in the automotive industry, ferritic steel sheets covered with aluminium alloy protective coatings are most often used presently. Aluminium coatings offer steel protection against the action of a corrosion medium at an elevated temperature, including combustion gas, and in the case of AlSi hot-dip coatings, also abrasion resistance [18, 19, 13]. Recently, it has been increasingly preferred to use austenitic steels in the production of exhaust systems [3, 4]. These steels have an extremely favourable combination of chemical properties and plastic forming capabilities [4, 6, 17]. The global production of these steels is still at a high level with a growing trend. 95% of the production of austenitic stainless steels are plastically shaped products, of which almost 10% are used in the automotive industry [6, 8]. The good plastic deformability of austenitic steels compensates for both their higher price, as well as the use of costly technologies, such as hydroforming.

Hydroforming is a method of forming flat metal sheets and closed sections using a fluid (most often water) under pressure (Fig. 1) [11, 1]. The advantage of the method in question is a reduction of the number of welded joints in structures and obtaining parts with a better surface condition, thinner walls and better dimensional tolerance [14, 15, 5]. At present, forming sections by this method is most commonly used in the manufacture of bicycles (aluminium frames) and in the automotive industry. By hydroforming, car bodies, carrying frames, mufflers and other parts, including exhaust system components, are produced [9].

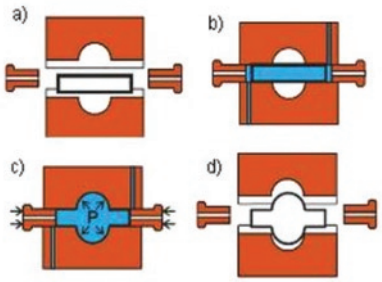


Fig. 1. Schematic diagram of the hydroforming operation: a – putting the pipe in the die, b – introducing a fluid inside the pipe, c – forming the pipe shape under fluid pressure, d – discharging the fluid and taking the formed part out of the die

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

By the hydroforming method it is possible to obtain complex shapes of parts with a varying curvature, which difficult to obtain by traditional plastic working method [3, 9, 10]. This is extremely important from the point of view of packing numerous items of car mechanics in the smallest space possible. At the same time, for closed sections, they must allow the free flow of media, such as combustion gas.

The specific conditions of hydromechanical forming of tubes, in which the material has no possibility of free "flowing" in the edge region, as is the case for sheets, cause high stresses to form in the material [7, 20]. The high stress level in a device's part makes it susceptible to dimensional instability. What is more, even small mechanical or corrosion damage initiated during the operation of such a part will generate its disproportionally large deformation due to stress relaxation. Therefore, from the point of view of its service life it is essential to determine that stress by non-destructive methods and then to mitigate it. The present paper reports the results of stress measurements in an exhaust system pipe made by bending technology and hydroforming technology, respectively.

2. Material and testing methodology

The subject of research was a pipe designated for exhaust system, in which its final shape was imparted in the manufacturing process by hydroforming technology (Fig. 2). The pipe subjected to hydroforming had a wall thickness of max. 1.7 mm and was made of chromium-nickel steel in grade *X5CrNi18-10* (AISI 304L) with an austenitic microstructure (Fig. 3). The nickel concentration in the steel, as determined spectroscopically, was 9.6 wt%, which imparts a deep draw quality (DDO) to the steel, compared to the standard version of grade 18-8.

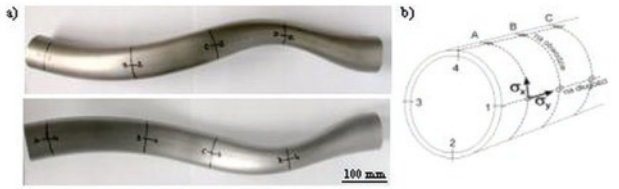


Fig. 2. a) A general view of the examined pipe and b) diagram of the location of the stress measurement places (σ) at outer surface of pipe. Sings: A, B, C and D – pipe perimeter, 1, 2, 3 and 4 – points on the pipe perimeter positioned every 90°

The objective of the research was to determine the stress on the outer surface of the pipe in the circumferential (x) and longitudinal (y) directions. For the purposes of testing, four regions (perimeters) were sectioned off on the pipe, of which three (denoted as B, C and D – Fig. 2a) were situated in the locations of a great change in pipe shape and one (denoted as A – Fig. 2a), near the pipe end, where the pipe cross-section was the closest in shape to circular The stress magnitudes were determined at four points on the pipe perimeter (denoted as 1, 2, 3 and 4 – Fig. 2b), positioned every ~ 90° in such a manner that points, e.g. A1, B1, C1 and D1, be situated along one pipe generating line.

For determining the stress values, an X-ray diffraction method, referred to as the $\sin^2\psi$ method, was employed [16, 2]. Tests were performed based on the diffraction reflection from plane (311), which is preferably used in stress measurements in austenitic steels [12]. The measurements were taken using a PROTO diffractometer dedicated for stress measurements at the Materials Science Department, Faculty of Machines Construction and Aircraft Engineering, of the Rzeszow University of Technology (Fig. 4).

K_αMn radiation (a ∅2mm collimator) of a wavelength of 0.2103 nm was used, which enabled stress to be measured in a sub-surface steel layer of a maximum thickness of 17 μm.

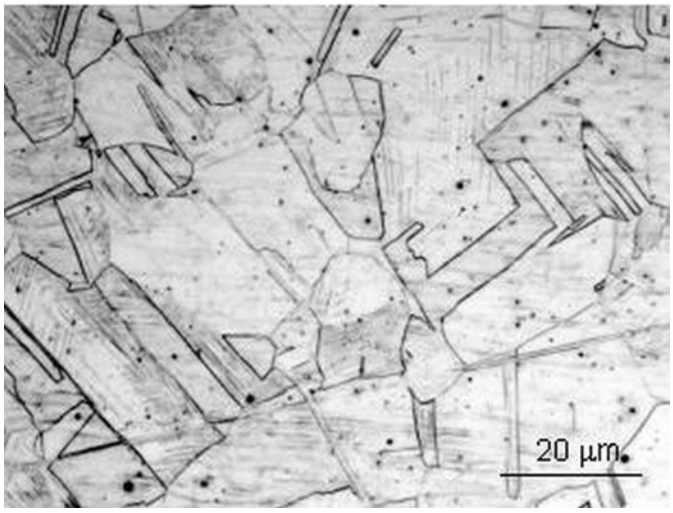


Fig. 3. Steel microstructure steel on the pipe cross-section

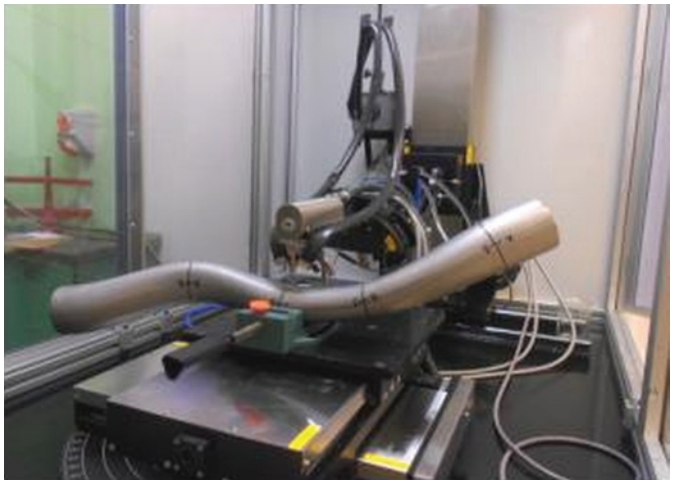


Fig. 4. The measurement of stress in the hydroformed pipe using a PROTO diffractometer

The stress determination by the X-ray method relies on the determination of crystal lattice deformation, ε, caused by, among other factors, by plastic working of polycrystalline material. This deformation is defined as the difference between interplanar distances Δ*d*, in the material with stress and without stress. The stress, σ_ϕ, is calculated from relationship (1), where ϕ defines the stress direction (selected by positioning the part during measurement), while ψ - angle of diffractometer head positioning or part surface inclination in the measurement of d_{hkl} of the deformed lattice:

$$\varepsilon_{\varphi\psi} = \Delta d / d_o = \left(\frac{1+\nu}{E}\right)\sigma_\varphi \sin^2\psi + \left(\frac{\nu}{E}\right)(\sigma_{11}+\sigma_{22}) \qquad (1)$$

where: d_o – distance between the lattice planes in the undeformed material ($d_o^{311}{}_{(aust)}$=0.1083 nm); σ_{11} and σ_{22} - principal stresses in the part surface (due to the measurement depth not exceeding a dozen or so μm, it is assumed that σ_{33}=0); ν - Poisson's ratio; E – Young's modulus [16].

The stress was calculated assuming the following X-ray elasticity constants for planes (311): ½ s_2=6.33×10^{-6} MPa^{-1} and $-s_1$=1.42×10^{-6} MPa^{-1} (the XRDWin software), whose values correspond to the mechanical constants of Young's modulus of E=200 GPa and the Poisson ratio of ν=0.29, according to relationship (2):

$$E = 1/(s_1 + \tfrac{1}{2}s_2) \quad \text{and} \quad \nu = -s_1/(s_1 + \tfrac{1}{2}s_2)\,. \qquad (2)$$

3. Results

The values of stresses determined by measurements on the pipe outer surface, as schematically shown in Figure 3b, are represented in Figures 5 and 6. All of the determined stresses were tensile stresses, both in the circumferential and longitudinal directions. The stress values were characterized by a large scatter, which was greater for longitudinal stress – the range of 21÷253 MPa, compared to the stress in the circumferential direction – the range of 65÷227 MPa.

The analysis of the stress distribution on the pipe perimeter (Fig. 5) showed that the greatest variation in stress magnitudes in both of the examined directions occurred on perimeter C in the central part of the pipe length. The most uniform stress distribution was found on perimeter D, whereas, circumferential stresses were, on average, greater that longitudinal stresses by approx. 80 MPa. The analysis of the stress distribution on the pipe length (Fig. 6) showed that the greatest diversity of circumferential stress magnitudes occurred on generating lines 4 and 3, while longitudinal stress magnitudes, on generating line 2. The most uniform distribution of circumferential stresses was found on generating lines 1 and 2, while longitudinal stresses, on generating line 3.

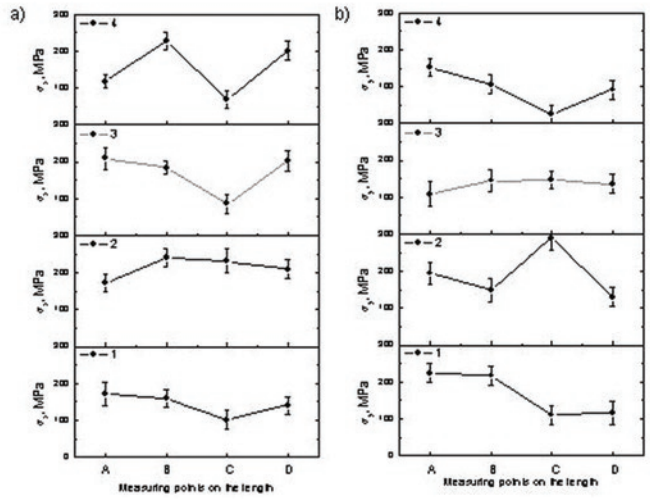


Fig. 5. The stress distribution on the pipe perimeter: a) circumferential stresses and b) longitudinal stresses

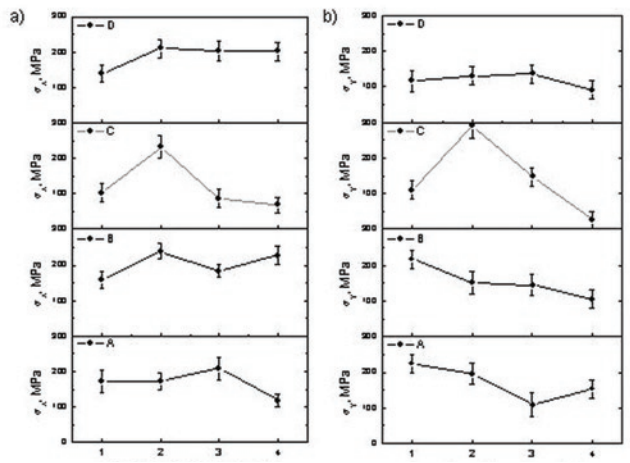


Fig. 6. The stress distribution on the pipe length a) circumferential stresses and b) longitudinal stresses

More generalized information on the distribution of residual stresses on the pipe surface can be provided by the averaged stress values for individual pipe regions (Fig. 7) and the coefficients of variation (Table 1) defined as (3):

$$V = S/\bar{\sigma} \qquad (3)$$

where: S – standard deviation , $\bar{\sigma}$ - arithmetic average.

In the analysis of the average values of $\bar{\sigma}_x$ and $\bar{\sigma}_y$ stresses on respective pipe perimeters (A, B, C and D), no relationship between them was found (Fig. 7a). It can only be noted that the circumferential stress $\bar{\sigma}_x$ was the lowest one the central part of the pipe (on perimeter C). The average stress in the longitudinal direction, $\bar{\sigma}_y$ (Fig. 7a), was the highest in the vicinity of the pipe end represented by perimeter A, and decreased steadily towards perimeter D. It was also found that both circumferential stress and longitudinal, as determined in region C, exhibited the greatest coefficients of variation, V (61% and 77%), which means a large scatter of stress values in this pipe region. In turn, in region D, stresses in both directions showed the smallest coefficients of variation (17%), which means that a small scatter of stress values occurred there.

A similar trend in their distribution was found in the analysis of the average $\bar{\sigma}_x$ and $\bar{\sigma}_y$ stress values on individual pipe generating lines (1, 2, 3 and 4) (Fig. 7b). The highest average stresses characterized generating line 2, while the lowest, generating line 4. These generating lines lay on the opposite pipe walls. A large scatter of stress magnitudes in both directions occurred along generating line 4 (V=48% and 56%). A small scatter was shown by circumferential stress on generating line 2, and by longitudinal stress, on generating line 3.

Notwithstanding the similarity in circumferential and longitudinal stress distributions on the pipe generating lines, it should be underlined that the analysis of longitudinal stress distribution is not often seen as particularly useful for pipes. This is due to the fact under the pressure of a medium inside the operated pipe, it is primarily the circumferential stress that becomes augmented. Its magnitude on the outer surface is two times greater than that of longitudinal (axial) stress. For this reason, the risk of a pipe being damaged is associated chiefly with the magnitudes of circumferential stress – and it is this stress that both manufacturers and customers recommend to be determined. A pipe in a technological condition was investigated in this study. Under liquid pressure during hydroforming, no free widening of the pipe results due to the restriction of its shape by the die, so it can be presumed that the longitudinal stress may be relatively higher compared to circumferential stress. The stress values determined experimentally confirm this presumption with respect to the outer pipe surface.

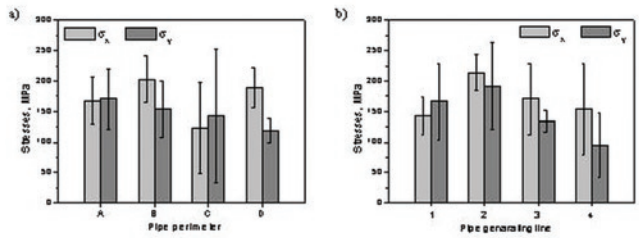


Fig. 7. The averaged stress values in respective regions of a hydroformed pipe: a) on the perimeters and b) on the generating lines (the error marks represent the standard deviation)

Due to the shape of the pipe and a small thickness of its wall it was not possible to take accurate hardness measurements directly on the

Table 1. The coefficients of variation, V, of stress in different hydroformed pipe regions

Circuit	$V(\sigma_x)$	$V(\sigma_y)$	Generating line	$V(\sigma_x)$	$V(\sigma_y)$
A	22	29	1	21	37
B	19	30	2	14	37
C	61	77	3	34	13
D	17	17	4	48	56

pipe surface, neither by the Vickers nor ultrasonic method. Nevertheless, the results of an attempt to take such measurements in approximation along generating line 1 are shown in Figure 8. The measurements were taken on the pipe surface between perimeters A, B, C and D in order not to damage the locations selected for stress measurements. Notwithstanding the only approximate character of the determined hardness values, it can be noticed that they reflected, to some extent, the distribution of σ_x and σ_y stresses along that generating line, i.e. they were smaller where stress values were also smaller. After stress measurements, sections were taken from the pipe to determine the wall thickness, which are also shown in Figure 8. The presumption that stresses are the highest in pipe regions with the smallest wall thickness, reduced during forming, was confirmed.

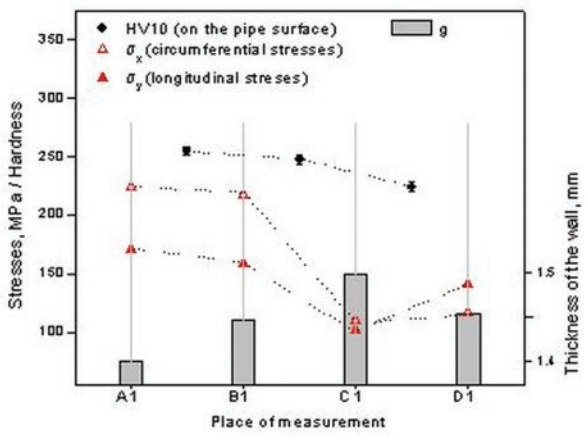


Fig. 8. Comparison of stress, hardness and wall thickness values on the pipe length (generating line 1)

4. Discussion and summary

A complex tensile stress state has been found on the outer surface of a chromium-nickel steel exhaust system pipe formed by hydroforming technology. The magnitudes of the highest surface stresses exceed the level of the yield point of steel 304L in version DDQ (approx. 170MPa). This means that the steel was strain hardened during hydroforming and, as can be presumed considering the complex shape of the pipe, also as a result of pre-bending preceding the hydroforming operation.

The stresses were characterized by a wide range of variability both in the circumferential (σ_x) and the longitudinal (σ_y) directions, amounting to 69÷240 MPa and 26÷290 MPa, respectively. The analysis of the average values of stresses and variation coefficients found that the lowest stresses, having simultaneously the greatest scatter, occurred in the central part of the pipe, where the reduction of the wall thickness was the least. Considering the surface character of the determined stresses, the source of that scatter should be sought in differences in the conditions of friction against the die between different pipe surface fragments during hydroforming.

The distribution of surface stresses determined by the diffractometric method is generally consistent with the computer-generated model of strain distribution in individual pipe regions, shown in Figure 9. It should be emphasized, however, that at the wall thicknesses (<1.5 mm) possessed by the examined pipe, the model represents rather average strains within the whole wall thickness. It will not reflect any possible incidental phenomena (associated e.g. with transport and storage) which could occur in production conditions.

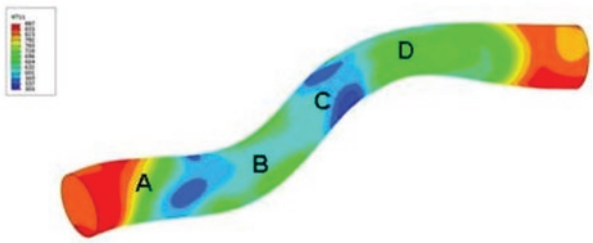


Fig. 9. The distribution of strains in the hydroformed pipe, generated from computer modelling

The values of determined stresses and the analysis of their distribution, based on measurements taken by the X-ray method, concern small surface areas defined by the cross-section of collimated radiation (∅2mm) and its penetration into the steel (approx. 17 μm) at the outer surface of the pipe. Depicting the distribution of stresses or strains within so thin layer is not achievable in computer modelling of plastic forming processes. Therefore, the X-ray method may only be used as a valuable complement to modelling, especially as modern diffractometers enable non-destructive measurements to be taken on thin-walled products of a complex surface shape.

Numerical analyses of the flow of a liquid medium in engine piping, presented in the literature, show that the medium exerts different pressure on the pipe walls in different pipe locations. This pressure depends primarily on the pipe bend angle and medium parameters, such as temperature, density or flow velocity [20]. This suggests that in locations, where the pipe curvature is the greatest, making stress distribution mapping by comprehensive measurements on the perimeters and along the pipe generating lines would be advisable.

References

1. Alaswad A, Benyounis K Y, Olabi A G. Tube hydroforming process: a reference guide. Materials and Design 2012; 33: 328-339, https://doi.org/10.1016/j.matdes.2011.07.052.
2. Baczmański A, Wierzbanowski K, Lipiński P. Determination of Residual Stresses in Plastically Deformed Polycrystalline Material. Materials Science Forum 1994; 157-162: 2051-2058, https://doi.org/10.4028/www.scientific.net/MSF.157-162.2051.
3. Bahman K. Trends for stainless steel tube in automotive applications. The Tube & Pipe Journal, September 13, 2005 (thefabricator.com).
4. Brytan Z. Stainless steel in the automotive industry (in Polish). STAL Metale & Nowe Technologie 2013; 11-12: 14-19.

5. Chałupczak J. Hydromechanical spreading in application to the formation of tees and X-pieces (in Polish). Works of the Kielce University of Technology. Mechanics; 39. Habilitation dissertation. Kielce, 1986.
6. Gronostajski Z, Kuziak R. Metallurgical, technological and functional foundations of advanced high-strength steels for the automotive industry (in Polish). Works of the Institute of Ferrous Metallurgy 2010; 22-26.
7. Hashemi R, Assemoiur A, Masourni E, Abad K. Implementation of the forming limit stress diagram to obtain suitable load path in tube hydroforming considering M-K model. Materials & Design 2009; 30(9): 3545-3553, https://doi.org/10.1016/j.matdes.2009.03.002.
8. ISSF International Stainless Steels Forum. Stainless Steel Consumption Forecast, October 2017, (http://www.worldstainless.org/statistics) 05.12.2018
9. Kocańda A, Sadłowska H. Automotive component development by means of hydroforming. Archives of Civil and Mechanical Engineering 2008; 8(3): 55-69, https://doi.org/10.1016/S1644-9665(12)60163-0.
10. Koç M. An overall review of tube hydroforming (THF) technology. Journal of Materials Processing Technology 2001; 108: 384-393, https://doi.org/10.1016/S0924-0136(00)00830-X.
11. Koç M (Ed.). Hydroforming for Advanced Manufacturing. Woodhead Publishing Limited England, and CRC Press USA, 2008, https://doi.org/10.1533/9781845694418.
12. Kucharska B, Krzywiecki M. Stresses in a Cr-Ni superficial steel layer based on x-ray measurements and electropolishing Solid State Phenomena 2015; 223: 348-354, https://doi.org/10.4028/www.scientific.net/SSP.223.355.
13. Kucharska B, Wróbel A, Kulej E, Nitkiewicz Z. The X-ray measurement of the thermal expansibility of Al-Si alloy in the form of cast and a protective coating on steel. Solid State Phenomena 2010; 163: 286-290, https://doi.org/10.4028/www.scientific.net/SSP.163.286.
14 Miłek T. Variations of wall thickness in the sections of hydromechanically bulged copper cross joints. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2003; 2(18): 45-48.
15. Morphy G. Pressure-sequence and high pressure hydroforming: Knowing the processes can mean boosting profits. The Tube & Pipe Journal, September/October 1998 (thefabricator.com, February 2001).
16. Skrzypek S J, Witkowska M, Kowalska M, Chruściel K. The non-destructive X-Ray methods in measuring of some material properties (in Polish). Hutnik-Wiadomości Hutnicze 2012; 79(4): 238-246.
17. Susceptibility of stainless steels to plastic working. Euro Inox, Series: Materials and applications. Book No. 8, 2008.
18. Wróbel-Knysak A, Kucharska B, The abrasion of Al-Si coatings with different silicon crystal morphology used in car exhaust systems. Tribologia 2016; 5: 209-218, https://doi.org/10.5604/01.3001.0010.6701.
19. Xianfeng Chen, Zhongqi Yu, Bo Hou, Shuhui Li, Zhongqin Lin. A theoretical and experimental study on forming limit diagram for a seamed tube hydroforming. Journal of Materials Processing Technology 2011; 211(12): 2012-2021, https://doi.org/10.1016/j.jmatprotec.2011.06.023.
20 Kumbár V, Votava J, Numerical modelling of pressure and velocity rates of flowing engine oils in real pipe. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 7(3): 422-426, https://doi.org/10.17531/ein.2015.3.13.

Barbara KUCHARSKA
Department of Production Engineering and Materials Technology
Czestochowa University of Technology
Armii Krajowej 19 str., 42-200 Czestochowa, Poland

E-mail: kucharska.barbara@wip.pcz.pl

Article citation info:
KANG R, WANG J, CHENG J, CHEN J, PANG Y. Intelligent forecasting of automatic train protection system failure rate in China high-speed railway. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 567–576, http://dx.doi.org/10.17531/ein.2019.4.5.

Renwei KANG
Junfeng WANG
Jianfeng CHENG
Jianqiu CHEN
Yanzhi PANG

INTELLIGENT FORECASTING OF AUTOMATIC TRAIN PROTECTION SYSTEM FAILURE RATE IN CHINA HIGH-SPEED RAILWAY

INTELIGENTNE PROGNOZOWANIE INTENSYWNOŚCI USZKODZEŃ AUTOMATYCZNEGO SYSTEMU OCHRONY POCIĄGÓW KOLEI DUŻYCH PRĘDKOŚCI W CHINACH

Intelligent and personalized dynamic maintenance and spare parts configuration of high-speed railway have been the main trend to guarantee the safety capability of trains. In this paper, a new Automatic Train Protection (ATP) system failure rate calculation method is proposed, and the delay time and embedded dimension are determined by C-C algorithm. Then the phase space is reconstructed from one-dimensional time series to high-dimensional space. Based on chaotic characteristics of failure rate, a short-term intelligent forecasting model of failure rate of ATP system is established. The actual failure statistics from 2010 to 2018 are used as samples to train and test the validity of the model. From prediction results, it shows that the proposed chaos prediction model has an accuracy of 99.71%, which is better than the support vector machine model. Through the intelligent prediction of failure rate, this paper solves the maintenance inflexibility and imbalance of supply and demand of spare parts configuration.

Keywords: *intelligent maintenance, high-speed railway, failure rate, automatic train protection system, prediction model, chaos.*

Inteligentna i spersonalizowana dynamiczna konserwacja i konfiguracja części zamiennych pociągów kolei dużych prędkości stanowią ostatnio główny trend w zakresie zapewniania bezpieczeństwa pociągów. W niniejszym artykule zaproponowano nową metodę obliczania intensywności uszkodzeń systemu Automatycznej Ochrony Pociągu (ATP), a czas opóźnienia i wymiar zanurzenia określano za pomocą algorytmu CC. Następnie, przestrzeń fazową przekształcono z jednowymiarowego szeregu czasowego do przestrzeni wielowymiarowej. Opierając się na chaotycznych charakterystykach intensywności uszkodzeń, utworzono model krótkoterminowego inteligentnego prognozowania awaryjności systemu ATP. Do uczenia modelu i weryfikacji jego trafności wykorzystano rzeczywiste dane statystyczne dotyczące awarii pociągów z lat 2010–2018. Z wyników prognoz wynika, że proponowany model predykcji, oparty na teorii chaosu, cechuje się dokładnością na poziomie 99,71%, czyli wyższą niż model maszyny wektorów nośnych. Dając możliwość inteligentnej predykcji intensywności uszkodzeń, niniejsza praca rozwiązuje problem braku elastyczności w utrzymaniu ruchu pociągów oraz braku równowagi między podażą a popytem na części zamienne.

Słowa kluczowe: *inteligentna konserwacja, kolej dużych prędkości, intensywność uszkodzeń, automatyczny system ochrony pociągu, model predykcyjny, chaos.*

1. Introduction

China high-speed railway has achieved remarkable achievements all over the world. On August 1 of 2008, the first China high-speed railway, Beijing-Tianjin intercity railway, with a speed of 350 km/h was operated. By the end of 2018, the operation mileage of high-speed railway is 29,000 kilometers, accounted for more than 2/3 of the total mileage all over the world, and has transported more than 9 billion passengers. China has become the world's high-speed railway country with the longest operation mileage, the highest transport density, and the most complex operating scenarios.

In 2018, there are 2811 high-speed trains all over China. The average number of operating high-speed trains per day are 2186, the average operation mileage per day is 3.88 million kilometers, and the largest number of passengers per day reach 8.10 million, ranking first in the world. The data shows that high-speed trains have become main means of transportation for 1.3 billion Chinese people. From 2008 to 2018, the growth trend of the operating mileage, the total number of trains, the average operating number and the mileage of trains per day are shown in Fig.1. As it can be seen from the figure, the scale of China high-speed railway is steadily increased year by year, it will continue to expand in the future.

China is the country with the longest high-speed railway and the largest safe transportation of high-speed railway. ATP (Automatic Train Protection) system is the core system to guarantee the safe operation of high-speed trains. In each train, two sets of ATP system are located at both ends of the train. Through calculating the train speed profile, ATP determines the running speed, the position and the headway towards the preceding train, so as to protect the train from

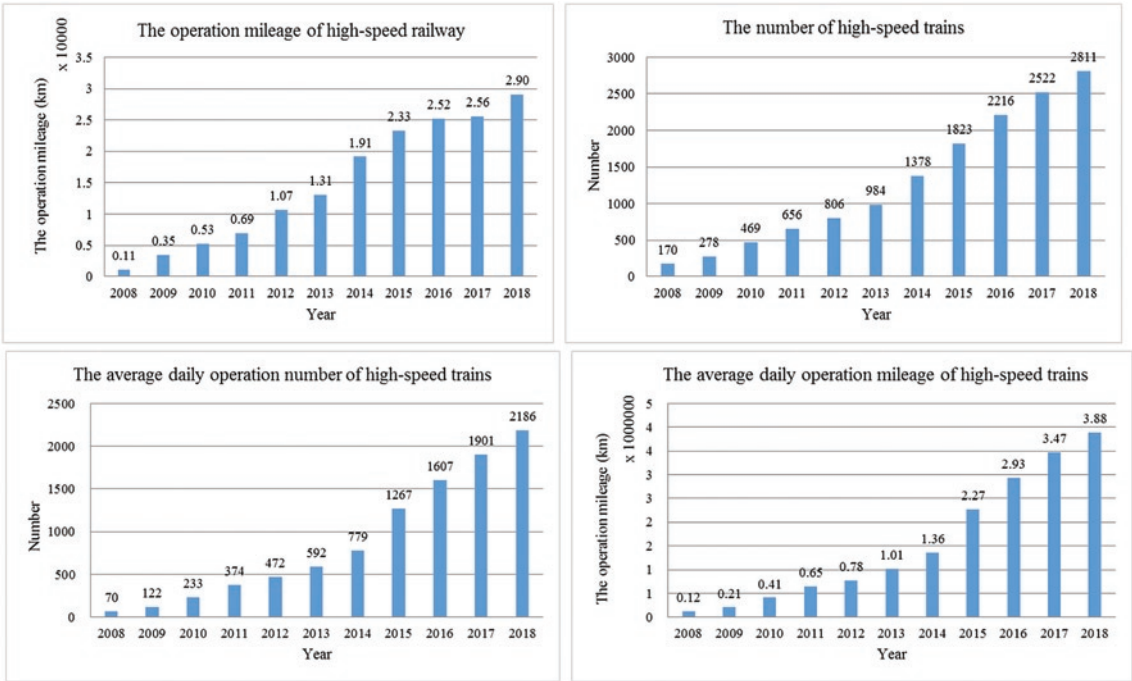


Fig.1. Growth trend of China high-speed railway scale

collision, derailment, and over-passing signals at danger. Therefore, the failure of ATP system will affect the safe operation of trains.

ATP system equipment includes VC (Vital Computer), SDP (Speed & Distance Processing unit), BTM (Balise Transmission Module), TCR (Track Circuit Reader), DMI (Driver-Machine Interface), TIU (Train Interface Unit), antenna, speed sensor, JRU (Juridical Recorder Unit), etc., where all devices cooperate with each other to guarantee the safe operation of train. As the scale of high-speed railway construction expands, the number of ATP systems is increasing. What's more, some ATP systems have been in operation for over 10 years, and a large number of ATP systems gradually enter the high-risk period over time. Therefore, better scientific maintenance strategies [11, 15] are urgently needed. However, there are still some deficiencies in application and maintenance of ATP systems.

Firstly, the maintenance strategy of ATP system is fixed and the maintenance mode is single. At present, China railway stipulates that ATP system should carry out planned maintenance in accordance with fixed maintenance intervals and items [5]. ATP system maintenance is divided into two different levels. The first is to maintain once a day on average, and the second is once a month on average. This mode will result in waste of maintenance resources or disrepair.

Secondly, the number of spare parts of ATP system are fixed. In order to replace and repair faulty equipment, China railway stipulates that each type of equipment of ATP system reserves spare parts in a fixed proportion [5], which influences maintenance owing to deficiencies and surplus of equipment. Since this mode can not balance the supply and demand well, it restricts the ability of emergency response.

Next, all kinds of ATP systems are maintained and configured in the same way, regardless of their failure rate. At present, there are five types of ATP systems in China, namely 300T, 300S, 300H, 200H and 200C. Different types of systems correspond to different production processes, design platforms, system structures, and software systems. Since their failure characteristics are different, so it should correspond to different maintenance strategies and reserve plans. However, there is only single way at present, which cannot dynamically adjust the maintenance strategy according to the equipment quality, or dynamically adjust the quantity of spare parts in a targeted manner. In such case, the equipment's safety capability is insufficient, and resources are not properly optimized.

Finally, the fixed maintenance strategy throughout the life cycle does not meet the inherent fault properties of the equipment. All ATP systems are maintained in the same strategy, regardless of service time, operation mileage, or natural environment. While, ATP equipment belong to electronic equipment, of which failure rate is a function of time, and the failure characteristics of the entire life cycle have obvious stages in line with the bathtub curve. The entire life cycle of ATP equipment can be divided into three phases: infant mortality, random failures, and wearout where each phase should correspond to elaborate maintenance strategy.

In order to solve the above problems, this paper proposes a method for intelligent forecasting failure rate of ATP system, which analyzes the characteristics of failure data and forecasts its short-term trend based on chaos from the perspective of history data. It can dynamically adjust the maintenance strategy and spare parts configuration in advance, and respond quickly to emergency situations, so as to prevent faults in advance and provide a strong guarantee for the safe operation of high-speed trains.

The rest of this paper is organized as follows. In section 2, the related work is reviewed. In section 3, the improved failure rate calculation method is proposed. In section 4, chaotic properties of ATP system failure rate series are determined. In section5, an intelligent prediction model is established. In section 6, the forecasting results are analyzed and applied. In section 7, the conclusion of our work is drawn.

2. Related work

The task of prediction is aimed at explicitly modeling variable dependencies and forecasting the next few values of a series [7]. On one hand, the failure rate of ATP system is affected by many factors such as service time, running mileage, natural environment, maintenance strategy, etc. So it is difficult to express influence mechanism with single functional relationship. On the other hand, ATP equipment belongs to the electronic device which is not separated from the common characteristics of the failure law, i.e., it contains inherent regular-

ity. Although the failure rate of ATP system shows great nonlinearity and seems to be irregular, it has a certain regularity.

Essentially, chaos is a nonlinear behavior existing between the realms of periodic and random [9]. Chaos theory extracts hidden regularity from the data directly to predict without subjective model [26]. Therefore, it is very suitable for analysis and research of ATP system failure rate.

The chaos theory itself has been studied. The research tries to sketch a theoretical view of chaos, i.e., to introduce chaotic and order properties, to compare them whenever possible, to find out how different a given order property and a given chaotic one are and to determine how convenient they are in the theory [2]. Besides, the research combines the Wiener chaos expansion approach to study the dynamics of a stochastic system with the classical problem of the prediction of a Gaussian process based on a realization of its past, to provide a general method for the construction [1].

On the other hand, chaos theory is used in the following aspects: analysis and prediction of short-term traffic flow [12, 29], railway freight traffic prediction [25, 31], big data analytic [23, 30] in intelligent transportation system and stock prediction [19]. It is used to analyze and predict certain relationships of transportation systems, which shows that in certain cases, fairly periodic systems analytically appear to be chaotic [9], and a soft computing approach and chaos theory have been successfully applied to solve the problem of early prediction of disruption in Joint European Torus machine [3]. The other study combines gradient descent method and chaos optimization method to predict leaching cycle period [18] and so on.

To summarize, the above researches provide a scientific and effective reference of data changes in the future. However, currently there is no research on failure rate forecasting of ATP system in China high-speed railway, and few literatures reported about intelligent maintenance from the perspective of prediction failure rate.

3. Improved failure rate calculation method

Failure rate in a quarter is used to measure the operation quality of different types of ATP systems in China railway. This evaluation index can explain the quality of systems to some extent. However, the failure of electronic devices has a cumulative effect. In other words, the device does not immediately fail when it is first put into use, and frequent failure is existing over time. Therefore, this paper uses the failure characteristics of electronic equipment, and adds cumulative effect factor to calculate the failure rate of ATP system based on existing calculation methods.

ATP system failure rate is one-dimensional time series:

$$x=\{x_i|i=1,2,\ldots,n\} \tag{1}$$

The existing failure rate is calculated as:

$$x_i=\frac{b_i}{r_i} \tag{2}$$

where x_i is the failure rate of a certain type of ATP system in the i-th quarter, b_i is the total number of faults of this type of ATP system in the i-th quarter all over the country, and r_i is the total operation mileage under the same conditions.

The improved failure rate is calculated as:

$$x_i=\frac{\sum_{j=1}^{i} a_j}{\sum_{j=1}^{i} s_j} \tag{3}$$

where x_i is the failure rate of a certain type of ATP system in the i-th day, a_j is the total number of faults of this type of ATP in the j-th day all over the country, and s_j is the total operation mileage on the same day and its unit is one million.

The existing method simply calculates the failure rate independently. Compared with it, the improved method proposed in this paper is obviously more scientific. It can be used as a key indicator to measure the quality of ATP system, but also to reconstruct the quality evaluation system fundamentally.

Firstly, it can be seen from equations (2) and (3) that with improved failure rate calculation method, time is refined and corrected from the original quarterly to daily. The calculation accuracy is higher.

Secondly, cumulative effects is added. The definition emphasizes the cumulative effect of ATP system over time, which reduces the fortuity of failure and better reflects the life cycle of the system.

What's more, since ATP system is powered on with the operation of high-speed trains, and powered off at the other time. Therefore, the accumulated operation time of ATP system is replaced by the accumulated operation mileage of high-speed trains, which can more accurately describe the characteristics of ATP system interval work.

4. ATP system failure rate series chaotic properties determination

Since one dimensional time series of ATP system failure rate is difficult to extract the potential failure regularity, so it should be mapped into high-dimensional space through phase space reconstruction theory, to determine whether it has chaotic properties. The SVM (Support Vector Machine) method [22] is used to train model in order to predict trends of ATP system failure in the short term. The whole idea is shown in Fig.2.

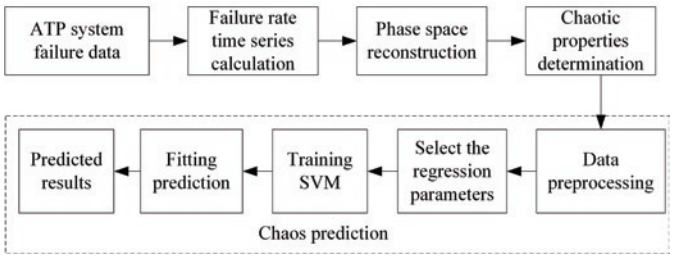


Fig. 2. The prediction flowchart of ATP system failure rate based on chaos

4.1. Original data collection and cleaning

Accurate collection of ATP failure data is the premise of failure rate prediction [16]. This paper obtains experimental sample data through the following steps.

1) Data collection and cleaning.

Firstly, collect fault conditions and operation mileage for each ATP system every day. The former includes time, location, summary, cause of the fault, and so on, and the latter includes start and end time, length of running, and so on. Then confirm that it is ATP system failure rather than the other systems failure through analysis and verification by professionals. i.e., the failure noise data is eliminated.

2) Determine the initial value.

In 2008, China high-speed railway started. In 2010, four types of ATP system, 300T-type, 300S-type, 200H-type, and 200C-type, were used on a large scale nationwide. Another 300H-type ATP system started to be used in large scale in 2013. Therefore, January 1 of 2010 is selected as the starting time point for failure data analysis. The data in the first three months after the specified starting time point is excluded since it has large fluctuations and error. Thus, the data from April 1, 2010 to December 31, 2018 is taken as a sample. The total

sample size is 3197, the training set sample size is 3047, and the test set sample size is 150. The sample size described below refers to the training set sample size, i.e., n=3047.

In units of days, xi is obtained in turn according to formula (3). The failure rate curve for each type of ATP system is shown in Fig.3.

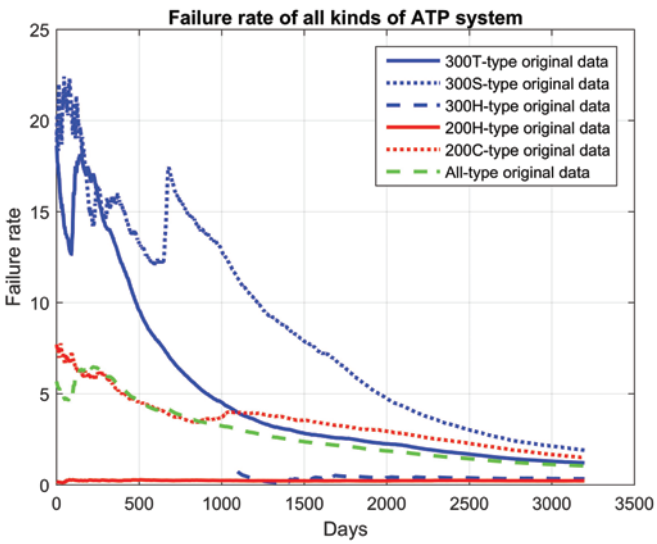


Fig. 3. Time series of all kinds of ATP system failure rate

In Fig.3, the ordinate represents the ATP system failure rate, and the green dotted line represents the failure rate for all types of ATP system. The following information can be obtained from Fig.3.

1) The failure rate shows overall downward trend and local fluctuations in a small range.

2) The system failure rates in 2010 and 2011 are significantly higher than in other years. The reasons are as follows: at the beginning of 2010, a variety of ATP systems have been used on a large scale for the first time, so the initial stability of system application is poor. In addition, China high-speed railway network system was still developing at that time, and its stability was relatively weak. Therefore, the failure rate is high. According to statistics, the number of ATP system failures in 2010 and 2011 accounted for 28.23% of the sum from 2010 to 2018, but the operation mileage only accounted for 7.36%, and the calculated failure rate was naturally slightly higher.

3) The 300H-type and 200H-type ATP systems are most stable with lowest failure rate.

In addition, it is very difficult to obtain other variation regularity of x. By extending one-dimensional time series [24] to the high-dimensional space, and guaranteeing the differential homeomorphism of the high-dimensional space and the original failure rate time series, the paper extracts the potential hidden regularity from high-dimensional space. Taken's embedding theorem [9, 31] shows that the geometrical features of ATP system failure rate time series are equal to the m-dimensional reconstruction space and have the same topology if there are appropriate delay time τ and embedded dimension m.

4.2. Phase space reconstruction

The one-dimensional time series x is reconstructed by delay embedded to obtain the trajectory matrix $\boldsymbol{Y}$:

$$Y=\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_{(m-1)\tau} \end{bmatrix}=\begin{bmatrix} x_1 & x_2 & \cdots & x_i & \cdots & x_{n-(m-1)\tau} \\ x_{1+\tau} & x_{2+\tau} & \cdots & x_{i+\tau} & \cdots & x_{n-(m-2)\tau} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{1+(m-1)\tau} & x_{2+(m-1)\tau} & \cdots & x_{i+(m-1)\tau} & \cdots & x_n \end{bmatrix} \quad (4)$$

where τ is delay time, and m is embedded dimension. Each column represents a phase point, and given $M=n-(m-1)\tau$, so $\boldsymbol{Y}$ is an $m\times M$ dimensional matrix.

The paper estimates delay time τ and embedded dimension m at the same time through C-C algorithm [20].

The time series x is subdivided into t disjoint time series [9] as following:

$$\{x(1), x(t+1), x(2t+1),\ldots\} \quad (5)$$

$$\{x(2), x(t+2), x(2t+2),\ldots\} \quad (6)$$

$$\{x(t), x(t+t), x(2t+t),\ldots\} \quad (7)$$

For each subseries, the statistic $S(m,n,r,\tau)$ is defined as following:

$$S(m,n,r,\tau)=\frac{1}{t}\sum_{l=1}^{t}\left\{C_l(m,n/t,r,\tau)-\left[C_l(1,n/t,r,\tau)\right]^n\right\} \quad (8)$$

where C_l is correlation integral of the l subseries, and its definition is:

$$C(m,n,r,\tau)=\frac{2}{M(M-1)}\sum_{1\le i<j\le M}\theta(r-\|X_i-X_j\|_\infty) \quad (9)$$

where r is neighbor radius, M is the number of phase points and $\theta(*)$ is Heaviside function.

Also, the distance of phase point X_i and the nearest phase point X_j is represented by infinite norm:

$$d=\|X_i-X_j\|_\infty=\max_{0\le k\le m-1}|x(i+k\tau)-x(j+k\tau)| \quad (10)$$

When $n\to\infty$, it can be written as:

$$S(m,r,\tau)=\frac{1}{t}\sum_{l=1}^{t}\left\{C_l(m,r,\tau)-\left[C_l(1,r,\tau)\right]^m\right\} \quad (11)$$

$S(m,r,\tau)\sim\tau$ reflect self-correlation of time series, and the paper takes improved C-C algorithm and defines:

$$\Delta S(m,\tau)=\mathrm{std}\{S(m,r,\tau)\} \quad (12)$$

where std{*} represents mean square deviation of $S(m,r,\tau)$.

The values of m and r are estimated properly according to statistical theory. For m=2,3,4,5, $r_i=i\sigma/2$, i=1,2,3. where σ is mean square deviation of time series [28]. Define:

$$S(t)=\frac{1}{16}\sum_{m=2}^{5}\sum_{j=1}^{4}S(m,r,t) \quad (13)$$

$$\Delta S(t)=\frac{1}{4}\sum_{m=2}^{5}\Delta S(m,t) \quad (14)$$

$$S_{cor}(t) = \Delta S(t) + |S(t)| \tag{15}$$

The optimal delay time τ corresponds to the first minimum point of $\Delta S(t)$ and the optimal embedding window τ_w corresponds to the global minimum point of the $S_{cor}(t)$. Because of embedding window [20] $\tau_w=(m-1)\tau$, we can obtain embedded dimension m.

By the end of December 2018, there were 2,811 high-speed trains and 5,622 ATP systems in China railway, ranking first in the world. Among which, with the number of 2110, 300T-type is the largest and most representative ATP system, accounted for 37.53%. Therefore, the 300T-type ATP system failure rate one-dimensional time series x is selected as the input, and the delay time and embedded dimension are calculated as shown in Fig.4.

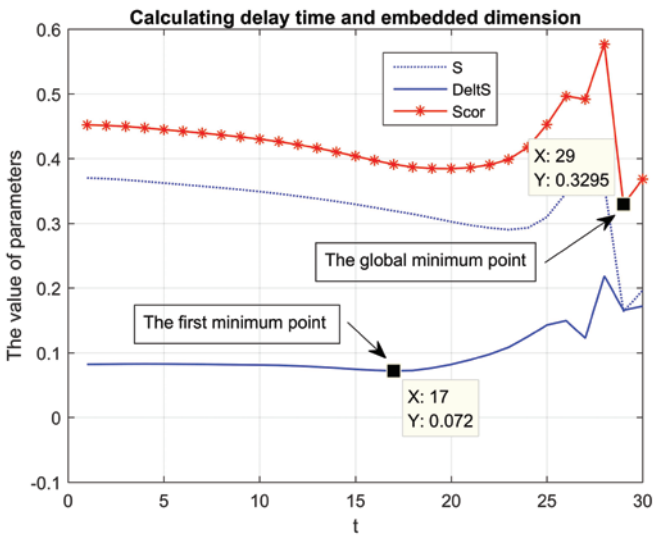


Fig. 4. Calculating delay time and embedded dimension by C-C algorithm

From Fig.4, the first minimum point of $\Delta S(t)$ is 17, so τ=17, and the global minimum point of the $S_{cor}(t)$ is 29, so τ_w=29, m=3.

Obtained delay time and embedded dimension, a size of 1×3047 one-dimensional time series x is reconstructed to 3×3013 high-dimensional matrix $\boldsymbol{Y}$.

4.3. Chaotic properties determination

It should be determined whether the series has chaotic properties after 300T-type ATP system failure rate time series x is reconstructed. The paper uses Lyapunov exponent method to judge chaotic properties. Grebogi [29] shows the system has chaotic properties as long as the maximum Lyapunov exponent λ is greater than 0.

The paper uses small data method to calculate the maximum Lyapunov exponent, and specific steps are as follows:

Step 1. Find the nearest neighbor $X_{\hat{j}}$ of phase point X_j in phase space, given:

$$d_j(0) = \min\left\|X_j - X_{\hat{j}}\right\|, \quad \left|j - \hat{j}\right| > P \tag{16}$$

where j=1,2,…M, and P represents the average cycle of time series. P is found by Fast Fourier Transformation, and P = 2.1145.

Step 2. Calculate the distance after phase point X_j evolution $i\Delta t$ time:

$$d_j(i) = \left|X_{j+i} - X_{\hat{j}+i}\right| \tag{17}$$

where i=1,2, …, min{M-j,M-$\hat{j}$}.

Step 3. Logarithm on both sides [29],

$$\ln d_j(i) \approx \ln d_j(0) + \lambda(i\Delta t) \tag{18}$$

The maximum Lyapunov exponent is the slope of the above straight line, it can be obtained approximately by the least squares method, i.e.:

$$y(i) = \frac{1}{q\Delta t}\sum_{j=1}^{q} \ln d_j(i) \tag{19}$$

where, q is the number of nonzero $d_j(i)$.

According to τ=17, m=3, the result of calculating the maximum Lyapunov exponent is shown in Fig.5.

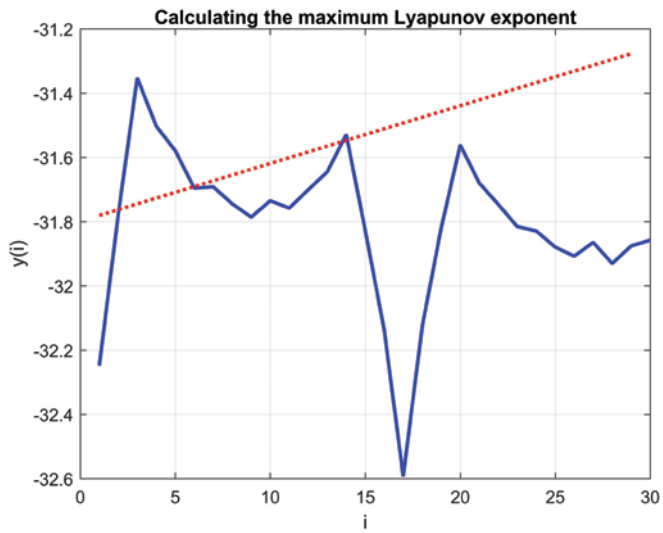


Fig. 5. Calculating the maximum Lyapunov exponent of 300T-type time series

In Fig.5, the solid line represents the value of $y(i)$, the dotted line indicates the line in step 3, whose slope is 0.018, namely the maximum Lyapunov exponent λ= 0.018 >0. Therefore, 300T-type failure rate time series has chaotic properties. It shows that the failure rate change of this ATP system follows certain rules, rather than irregular. The chaotic characteristics indicate that short-term predictions of failure rate can be made.

According to reference [25], the reciprocal of the maximum Lyapunov exponent (i.e., T=1/λ) represents the longest forecasting time of chaotic system. Because of λ=0.018, so T=1/λ≈56. It means that failure rate future 56 days can be predicated through 300T-type historical data. Therefore, it provides ample preparation time for the dynamic adjustment of maintenance strategy and spare parts configuration in advance. It also indicates that the intelligent prediction method proposed in this paper has the feasibility of implementation.

5. Intelligent prediction

After reconstructing the phase space and identifying the chaotic characteristics, the paper extracts potential failure regularity in a high-dimensional space, further to predict the trends and range of ATP system failure rate.

5.1. Prediction principle

300T-type failure rate time series $x=\{x_i|i=1,2,\ldots,n\}$ is made phase space reconstruction, and then it is obtained M phases point:

$$X(1)=\{x(1),x(1+\tau),\ldots, x(1+(m-1)\tau)\} \tag{20}$$

$$X(2)=\{x(2),x(2+\tau),\ldots, x(2+(m-1)\tau)\} \tag{21}$$

$$\ldots$$

$$X(M)=\{x(M),x(M+\tau),\ldots, x(M+(m-1)\tau)\} \tag{22}$$

Further evolution of phase points is:

$$X(M+1)=\{x(M+1),x(M+1+\tau),\ldots, x(M+1+(m-1)\tau)\} \tag{23}$$

The next sequence predict point $x(n+1)$ is exactly the element $x(M+1+(m-1)\tau)$.

There exists some definite and complex function connection between phase points $X(i)$ and $x(i+1+(m-1)\tau)$, namely:

$$x(i+1+(m-1)\tau)=f[X(i)] \tag{24}$$

Find an optimal function to fit $f[X(M_i)]$ with the SVM approximation ability [10], namely:

$$f[X(M_i)] = x(M_i+1+(m-1)\tau) \tag{25}$$

If it is determined $f(x)$, there:

$$x(n+1)=x(M+1+(m-1)\tau)=f[X(m)] \tag{26}$$

Similarly, it can forecast $x(n+2)$, $x(n+3)\ldots$

SVM is an effective machine learning method based on statistical learning theory, and has excellent learning performance and strong generalization ability [17, 27]. This paper needs to determine the function $f(x)$, which is clearly a regression problem. ε-support vector regression(ε-SVR) belongs to SVM and its goal of training is:

$$|y-f(x)|\leq\varepsilon \tag{27}$$

If specify the insensitivity ε, then get ε-SVR, i.e.:

$$\min_{w,b,\xi}\left[\frac{1}{2}\|w\|_2^2+c\sum_{i=1}^{l}\left(\xi_i+\xi_i^*\right)\right] \tag{28}$$

$$s.t.\quad y_i-\left\{\left[w,\Phi(x_i)\right]-b\right\}\leq\varepsilon+\xi_i \tag{29}$$

$$y_i-\left\{\left[w,\Phi(x_i)\right]-b\right\}\geq-\varepsilon-\xi_i^* \tag{30}$$

$$\xi_i\geq 0,\xi_i^*\geq 0, i=1,2,\cdots,l \tag{31}$$

Its solution is:

$$f(x)=\sum_{i=1}^{l}(a_i-a_i^*)K(x,x_i)-b \tag{32}$$

where c is the penalty coefficient and other parameters are available from references [13, 21]. So far, the model of intelligent forecasting of ATP system failure rate by chaos is shown in Fig.6.

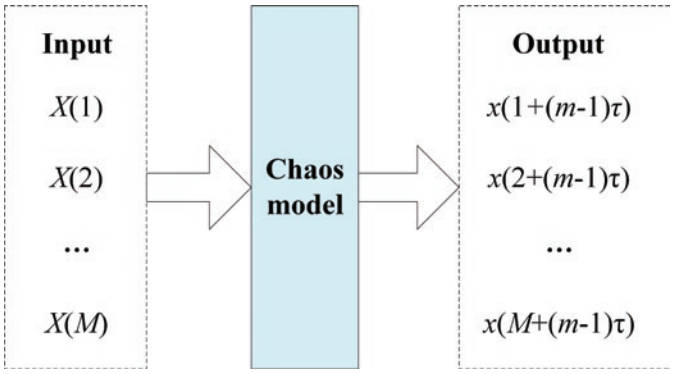


Fig.6. The model of intelligent forecasting of ATP system failure rate

5.2. Data normalization process

300T-type ATP system failure rate time series is reconstructed to $m\times M$ dimensional matrix $\boldsymbol{Y}$, i.e., the size of training set is 3×3013.

Select $(X(1),X(2),\ldots,X(M))$ as training set, that is X_train=$(X(1),X(2),\ldots,X(3013))$.

Select $(X(1),X(2),\ldots,X(M-1))$ as training set feature, that is X_train_ feature =$(X(1),X(2),\ldots,X(3012))$.

Select $(x(2+(m-1)\tau), x(3+(m-1)\tau), \ldots,x(n))$ as training set label, that is X_train_ label =$(x(36), x(37), \ldots,x(3047))$.

Early tentatively forecasting shows that the accuracy is low regarded matrix $\boldsymbol{Y}$ as an input directly without making data processing. Thus, training set feature and label are made normalization preprocessing according to the following map:

$$f: x\rightarrow y=\frac{x-\min(x)}{\max(x)-\min(x)} \tag{33}$$

In order to improve prediction accuracy, the element of matrix $\boldsymbol{Y}$ as chaos model input is structured into the range [0,1].

5.3. Model optimal parameters selection

The advantages and disadvantages of each parameter should be comprehensively considered to train the prediction model. To complement and balance the multiple parameters, optimal model is achieved. The training model parameter evaluation indicators are as follows:

Mean Square Error (MSE):

$$MSE=\frac{1}{n}\sum_{i=1}^{n}\left[f(x_i)-y_i\right]^2 \tag{34}$$

Squared Correlation Coefficient (r^2):

$$r^2=\frac{\left[n\sum_{i=1}^{n}f(x_i)y_i-\sum_{i=1}^{n}f(x_i)\sum_{i=1}^{n}y_i\right]^2}{\left(n\sum_{i=1}^{n}f(x_i)^2-\left(\sum_{i=1}^{n}f(x_i)\right)^2\right)\left(n\sum_{i=1}^{n}y_i^2-\left(\sum_{i=1}^{n}y_i\right)^2\right)} \tag{35}$$

where n represents the number of samples.

It needs to determine the three parameters [4, 8] for solving regression problems of $f(x)$ using SVR, which are insensitivity ε, penalty coefficient c and kernel parameter g.

The paper selects radial basis function:

$$K(x,y)=\exp\left(-g\left\|x-y\right\|^2\right) \tag{36}$$

where g is kernel parameter [6, 14].

First of all, set insensitivity ε=0.01. Then find the best c and g within the scope of the index of 2 by a cross validation method.

Roughly determine the scope of optimization c and g, and the values c are $2^{-6},2^{-5},\ldots,2^{6}$, and the values g are $2^{-6},2^{-5},\ldots,2^{0}$. The rough selection results are shown in Fig.7. Where the best c=0.25, g=1, and the minimum *MSE*=0.0038507.

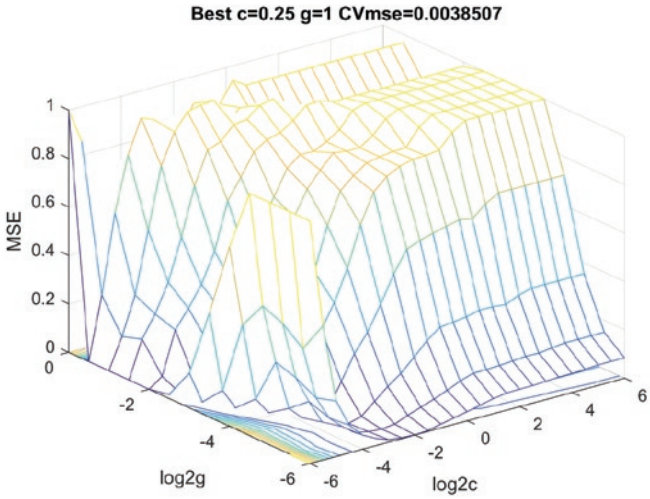


Fig. 7. Rough selection results of parameters c and g

In Fig.7, x-axis represents the value of 2 of c power, and y-axis represents the value of 2 of g, while z-axis represents mean square error under corresponding c and g.

In order to find a better parameter value, the value range of the penalty coefficient c is reduced to 2^{-6}~2^{-2}, and the range of the kernel parameter g is expanded to 2^{-6}~2^{2}. Thus, fine selection results of parameters c and g are shown in Fig.8. Where the best c=0.125, g=2, and the minimum *MSE*=0.003488. The minimum *MSE,* reduced by 9.42%, is less than the one by rough selection, indicating that the parameter results by fine selection are better.

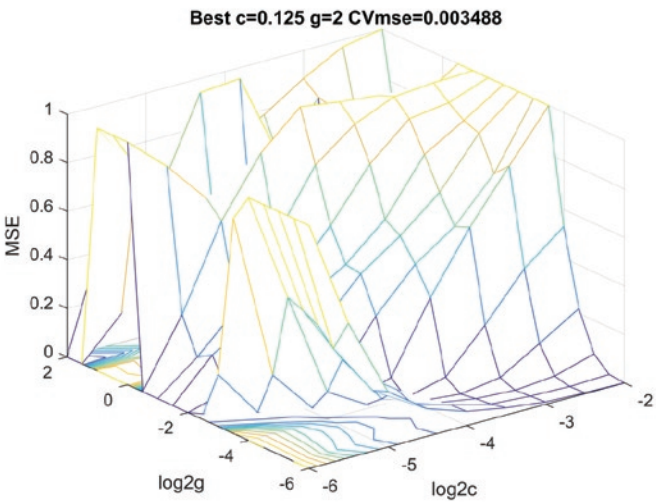


Fig. 8. Fine selection results of parameters c and g

5.4. Forecasting results

To train SVR according to the best parameters c and g in Fig.8, a regression model to fit $f(x)$ can be obtained based on the input-output relationship of Fig.6. Then the trend of future failure rate of 300T-type ATP system is predicted. At last, to set the predicted step length as 100, the model accuracy is verified. Fig.9 is a comparison between chaos forecasting data and the original data.

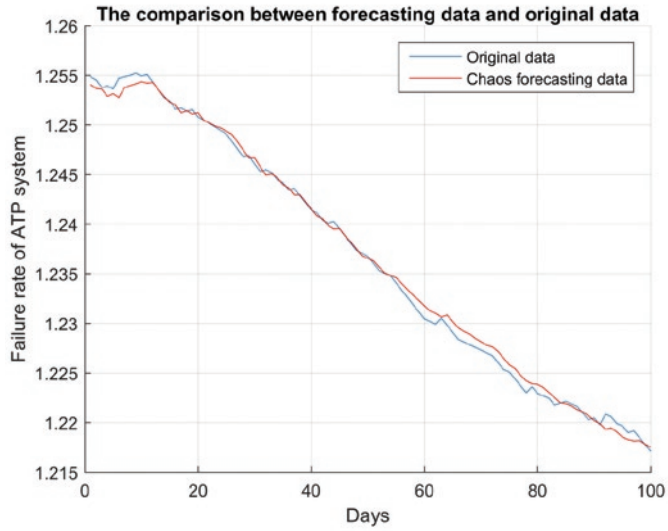


Fig. 9. The comparison between chaos forecasting data and the original data

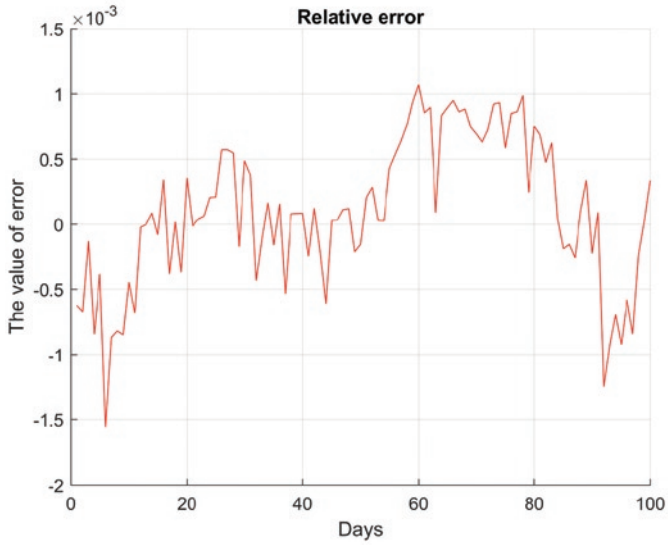


Fig. 10. Relative error of chaos prediction

In Fig.9, the data of chaos prediction and the original data are consistent, Mean Square Error *MSE*=0.0003416, Squared Correlation Coefficient r^2=0.997092, that is, the prediction accuracy reaches 99.71%. The relative error of chaos prediction is shown in Fig.10. The

error is relatively concentrated in the interval [-0.001, 0.001], which indicates that the chaos method is suitable for the failure rate trend prediction of ATP system in China high-speed railway.

6. Forecasting results analysis

This section analyzes the effectiveness of the proposed intelligent prediction method of failure rate from two aspects: model prediction accuracy and prediction results application.

6.1. Model prediction accuracy analysis

To illustrate the effectiveness of chaos intelligent forecasting on ATP system failure rate, this paper compares chaos forecasting results with the results of SVR method. The main differences between two methods are: chaos prediction makes the matrix $\boldsymbol{Y}$ as an input after phase space reconstruction, while SVR prediction makes one-dimensional time series x as an input directly; in addition, the former processing objects are phase points of high- dimensional matrix, while the latter are elements of one-dimensional time series. The main differences between the two methods are shown in Table 1.

The SVR prediction model still takes the 300T-type ATP system failure rate time series x as a sample, taking ε=0.01, and the 100 data as the test sets. Select the best parameter c=0.0442, g=5.6569, correspondingly, the minimum MSE=0.0041.

The mean squared error and squared correlation coefficient of chaos and SVR model are respectively shown in Table 2.

It can be seen from Table 2 that for the same 300T-type failure rate time series, chaos prediction mean square error is smaller than SVR prediction, while the former squared correlation coefficient is bigger than the latter, showing that chaos prediction accuracy is higher than the SVR.

In addition, the equipment failure rate is closely related to the quality of product, and the original quality determines the level of failure rate to a large extent, which coincides with the essential characteristics of the chaotic system's sensitivity to the initial value. It shows that the experimental results are consistent with the actual situation.

In summary, compared with SVR prediction, the prediction of the failure rate of ATP system based on chaos theory is more suitable and more accurate.

6.2. Analysis of the forecasting results application

Maintenance strategies and configure spare parts can be dynamically adjusted based on the volatility of forecasting results.

If the predicted failure rate is in upward trend, the frequency of maintenance should be increased, the maintenance items should be extended, and the equipment with higher failure rate should be overhauled to prevent the failure. At the same time, the spare parts of system should be adequately configured in advance to improve the emergency response capacity.

If the predicted failure rate is in a downward trend, the frequency of maintenance can be appropriately decreased, and the maintenance items can be simplified, which can reduce the workload of maintenance and alleviate the contradiction between insufficient and excessive reserves.

Therefore, the intelligence of the failure rate prediction method proposed in this paper is reflected in the following aspects:

1) For different types of ATP systems, to dynamically adjust different targeted maintenance strategies and spare parts reserve plans.

2) For a specific type of ATP system, to dynamically adjust maintenance intervals and items.

3) For a specific type of ATP system, to dynamically configure the quantity of spare parts.

4) For a set of ATP system, maintenance plans throughout the life cycle are dynamically adjusted based on factors such as service time, operation mileage, and natural environment and so on.

In summary, the differences between before and after using the intelligent prediction method presented in this paper are shown in Table 3.

Table 1. Difference between chaos model and SVR model

Difference point	Chaos model	SVR model
Processing objects	High- dimensional matrix $\boldsymbol{Y}$	One-dimensional time series x
Training set feature	$X(1),X(2),\ldots,X(M\text{-}1)$	$x(1),x(2),\ldots,x(n)$
Training set label	$x(2+(m\text{-}1)\tau), x(3+(m\text{-}1)\tau), \ldots,x(n)$	Day number sequence,1: n
One-step prediction	$X(M+1)$	$x(n+1)$

Table 2. Forecasting results comparison of chaos and SVR

Difference point	Chaos model	SVR model
MSE	0.0003416	0.0984576
r^2	0.997092	0.877012

Table 3. Using intelligent prediction methods before and after comparison

Difference point	Before the novel method	After the novel method
Maintenance strategy	Fixed maintenance cycle and content.	Dynamic maintenance cycle and content.
Spare parts reserve	The quantity is fixed.	The quantity is dynamic.
Maintenance strategy for different types of ATP systems	The same way.	The different way.
Spare parts reserve method for different types of ATP systems	The same way.	The different way.

7. Conclusion

On the basis of existing quality evaluation methods, this paper added the cumulative effect of electronic equipment over time, and improved the quality evaluation system of ATP system in China high-speed railway. By calculating the maximum Lyapunov exponent, it was revealed that failure rate change of ATP system had chaotic characteristics, in which it proved that ATP system failure rate time series was extremely sensitive to the initial conditions. Then a short-term intelligent prediction model of failure rate based on chaos was established, and the real failure data from 2010 to 2018 was selected as a sample for simulation experiments. The results showed that the accuracy of the intelligent prediction model was 99.71%, much higher than SVR prediction model.

According to the results, the dynamic adjustment to maintenance strategy and dynamic reserve spare parts were realized, and the personalized maintenance strategy of each ATP system through full life cycle was established. What's more, it improved the existing single and fixed maintenance mode of China railway, and solved the pain points such as rough maintenance mode, unbalanced supply and demand of spare parts, and insufficient emergency response capability. Finally, it realized the intelligent maintenance of ATP system in China high-speed railway.

As future work, the authors propose to explore the following research lines: (1) To further analyze the complex relationship among the failure rate, the natural environment, maintenance strategy and operation mileage and so on. (2) If there are more detailed classification of spare parts data, the method can make particular forecasting, to provide a reasonable reserves and maintenance strategy for a special spare parts.

Acknowledgements

This work was supported by State Key Laboratory of Rail Traffic Control and Safety under Grants RCS2018ZT010, by China Academy of Railway Sciences under Grants 2017YJ056, by Nanning special funds for small and high talents under Grants 2017042 and 2017043, by upgrading of basic ability of young and middle-aged teachers in Colleges and universities in Guangxi under Grants 2018KY0747 and 2018KY0748. The authors would like to thank Miss Junli Li, she revised grammatical and syntactical errors in the manuscript.

References

1. Alpay D, Kipnis A. Wiener Chaos Approach to Optimal Prediction. Numerical Functional Analysis and Optimization 2015; 36(10): 1286-1306, https://doi.org/10.1080/01630563.2015.1065273.
2. Blanchard F. Topological chaos: what may this mean. Journal of Difference Equations and Applications 2009; 15(1): 23-46, https://doi.org/10.1080/10236190802385355.
3. Cacciola M, Costantino D, Morabito F-C, Versaci M. Soft Computing and Chaos Theory for Disruption Prediction in Tokamak Reactors. International Journal of Modelling and Simulation 2008; 28(2): 165-173, https://doi.org/10.1080/02286203.2008.11442464.
4. Cao S-C, et al. Establishing a Flight Load Parameter Identification Model with Support Vector Machine Regression. Journal of Northwestern Polytechnical University 2013; 31(4): 535-539.
5. China railway standard. CTCS-2/3 Level Train Control on-board Equipment Maintenance Management Measures 2015; tiezongyun 57.
6. Dai A-N, et al. Intelligent control of a grain drying system using a GA-SVM-IMPC controller. Drying Technology 2018; 36(12): 1413-1435, https://doi.org/10.1080/07373937.2017.1407938.
7. Esling P, Agon C. Time-Series Data Mining. ACM Computing Surveys 2012; 45(1): 12-34, https://doi.org/10.1145/2379776.2379788.
8. Feng X-X, et al. Adaptive Multi-Kernel SVM With Spatial-Temporal Correlation for Short-Term Traffic Flow Prediction. IEEE Transactions on Intelligent Transportation Systems (Early Access) 2018: 1-13.
9. Frazier C, Kockelman M. Chaos Theory and Transportation Systems. Journal of the Transportation Research Board 2004; 1897: 9-17, https://doi.org/10.3141/1897-02.
10. Fu G, et al. Short-term Traffic Flow Forecasting Model Based on Support Vector Machine Regression. Journal of South China University of Technology(Natural Science Edition) 2013; 41(9): 71-76.
11. Galar D, Gustafson A, Tormos B, Berges L. Maintenance Decision Making based on different types of data fusion. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14 (2): 135-144.
12. Ghosh B, Basu B, O'Mahony M. Multivariate Short-Term Traffic Flow Forecasting Using Time-Series Analysis. IEEE Transactions on Intelligent Transportation Systems 2009; 10(2): 246-254, https://doi.org/10.1109/TITS.2009.2021448.
13. Guo Y-M, Ran C-B, Li X-L, Ma J-Z, Zhang L. Weighted prediction method with multiple time series using multi-kernel least squares support vector regression. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15 (2): 188-194.
14. Guo Y-M, Wang X-T, Liu C, Zheng Y-F, Cai X-B. Electronic system fault diagnosis with optimized multi-kernel SVM by improved CPSO. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16 (1): 85-91.
15. Jimenez Cortadi A, Irigoien I, Boto F, Sierra B, Suarez A, Ga lar D. A statistical data-based approach to instability detection and wear prediction in radial turning processes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (3): 405-412, https://doi.org/10.17531/ein.2018.3.8.
16. Kozielski M, Sikora M, Wróbel Ł. Decision support and maintenance system for natural hazards, processes and equipment monitoring. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (2): 218-228, https://doi.org/10.17531/ein.2016.2.9.
17. Liu B-L, et al. An improved PSO-SVM model for online recognition defects in eddy current testing. Nondestructive Testing and Evaluation 2013; 28(4): 367-385, https://doi.org/10.1080/10589759.2013.823608.
18. Liu C, Wu A-X, Yin S-H, Chen-X. Nonlinear chaotic characteristic in leaching process and prediction of leaching cycle period. Journal of Central South University 2016; 23(16): 2935-2940, https://doi.org/10.1007/s11771-016-3357-9.
19. Ma J-H, Qi E-S, Mo X. Application Study on Reconstruction of Chaotic Time Series and Prediction of Shanghai Stock Index. Systems Engineering-Theory & Practice 2013; 23(12): 86-93.
20. Meng Y-Y, Lu J-P, Wang J. Wind Power Chaos Prediction Based on Volterra Adaptive Filter. Power System Protection and Control 2012; 40(4): 90-95.

21. Nicolas P-C, Theodore B-T. On-line SVM learning via an incremental primal-dual technique. Optimization Methods & Software 2013; 28(2): 256-275, https://doi.org/10.1080/10556788.2011.633705.
22. Qu X, Wang W, Wang W-F, Liu P. Real-time rear-end crash potential prediction on freeways. Journal of Central South University 2017; 24(11): 2664-2673, https://doi.org/10.1007/s11771-017-3679-2.
23. Świderski A, Jóźwiak A, Jachimowski R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (2): 292-299, https://doi.org/10.17531/ein.2018.2.16.
24. Vališ d, Koucky M, Zak L. On approaches for non-direct determination of system deterioration. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14 (1): 33-41.
25. Wu H-W, Wang F-Z. Research on Railway Freight Traffic Prediction Based on Maximum Lyapunov Exponent. Journal of the China Railway Soci 2014; 36(4): 8-13.
26. Wu X-X. Li-Yorke chaos of translation semigroups. Journal of Difference Equations and Applications 2014; 20(1): 49-57, https://doi.org/10.1080/10236198.2013.809712.
27. Yan Z-G, Yao K, Yang Y-X. A novel adaptive differential evolution SVM model for predicting coal and gas outbursts. Journal of Difference Equations and Applications 2017; 23(1-2): 238-248, https://doi.org/10.1080/10236198.2016.1214725.
28. Zhang H-B, Sun X-D, He Y-L. Analysis and Prediction of Complex Dynamical Characteristics of Short-term Traffic Flow. Acta Physica Sinica 2014; 63(4): 1-8.
29. Zhang Y-M, Wu X-J, Bai S L. Chaotic Characteristic Analysis for Traffic Flow Series and DFPSOVF Prediction Model. Acta Physica Sinica 2013; 62(19): 1-9.
30. Zhu L, Yu F-R, Wang Y-G, et al. Big Data Analytics in Intelligent Transportation Syetems: A Survey. IEEE Transactions on Intelligent Transportation Systems 2019; 20(1):383-398, https://doi.org/10.1109/TITS.2018.2815678.
31. Zhu Z-H, Weng Z-S. Railway Passenger and Freight Volume Forecasting Based on Chaos Theory. Journal of the China Railway Soci 2011; 33(6): 1-7.

Renwei KANG
Junfeng WANG
State Key Laboratory of Rail Traffic Control and Safety
Beijing Jiaotong University
Haidian, Beijing 100044, China

Jianfeng CHENG
Signal & Communication Research Institute
China Academy of Railway Sciences
Haidian, Beijing 100081, China

Jianqiu CHEN
Yanzhi PANG
Rail Transit Scientific Research Institution
Nanning University
Nanning 530200, Guangxi, China

E-mails: krwjob@126.com, w2881@163.com, 583735920@qq.com, chenjq@bjtu.edu.cn, pangyanzhi@126.com

Article citation info:
JAMROZIAK K, KWASNIOWSKI S, KOSOBUDZKI M, ZAJAC P. Analysis of heat exchange in the powertrain of a road vehicle with a retarder. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 577–584, http://dx.doi.org/10.17531/ein.2019.4.6.

Krzysztof JAMROZIAK
Stanislaw KWASNIOWSKI
Mariusz KOSOBUDZKI
Pawel ZAJAC

ANALYSIS OF HEAT EXCHANGE IN THE POWERTRAIN OF A ROAD VEHICLE WITH A RETARDER

ANALIZA WYMIANY CIEPŁA W UKŁADZIE NAPĘDOWYM POJAZDU DROGOWEGO Z RETARDEREM

The paper presents a heat exchange model for the cooling system of any complex, physical system. Verification of the correctness of the theoretical model was carried out on the example of a vehicle with a combustion engine and additionally equipped with a hydraulic retarder. The results of laboratory tests, which were carried out on an engine test bench, were also performed for the above mentioned powertrain, so as to compare the results of modelling with the results of the tests. Determining the operating parameters of the components of the cooling system aimed at protecting the entire powertrain against overheating is a key task. Theoretical analysis of heat exchange in the powertrain of a road vehicle was carried out, with particular emphasis on the hydraulic retarder (a device braking the vehicle during a descent on roads with a high gradient of the road, mandatory according to the ADR convention). The subject of the study was a mathematical model of a complex cooling system developed by the authors, described by means of balance equations and differential equations. This model was tested with the use of the Matlab-Simulink suite for given load parameters of the cooling system, which were used in tests on an engine test bench. The values of coefficients describing the thermal state of the powertrain were obtained. Simulations were performed for different variants of technical parameters of the expanded cooling system. In this way, individual units and components of the cooling system were optimized so that it fulfilled its role in the assumed operating conditions and the ecologization of emission of energy sources (fuel) and harmful substances.

Keywords: *heavy-duty vehicle operation, vehicle retarder, brake support, heat exchange.*

Praca zawiera model wymiany ciepła układu chłodzenia dla złożonego, dowolnego układu fizycznego. Weryfikację poprawności modelu teoretycznego przeprowadzono na przykładzie pojazdu z silnikiem spalinowym oraz dodatkowo wyposażonym w retarder hydrauliczny. Wyniki badań laboratoryjnych, które przeprowadzono na hamowni również wykonano dla w/w zespołu napędowego, tak aby porównać wyniki modelowania z wynikami badań. Ustalenie parametrów eksploatacji elementów układu chodzenia, którego celem jest zabezpieczenie całego układu napędowego przed przegrzaniem to kluczowe zadanie. Przeprowadzono teoretyczną analizę wymiany ciepła w układzie napędowym pojazdu drogowego ze szczególnym uwzględnieniem zwalniacza hydraulicznego (urządzenia hamującego pojazd podczas zjazdu na drogach o dużym pochyleniu jezdni, obowiązkowe wg konwencji ADR). Przedmiotem badań był opracowany przez autorów model matematyczny rozbudowanego układu chłodzenia opisany za pomocą równań bilansowych i równań różniczkowych. Model ten testowano z wykorzystaniem pakietu Matlab-Simulink dla zadanych parametrów obciążenia układu chłodzenia, które wykorzystywano w badaniach w stanowiskowych na hamowni silnikowej. Uzyskano wartości współczynników opisujących stan cieplny jednostki napędowej. Symulacje wykonano dla różnych wariantów parametrów technicznych rozbudowanego układu chłodzenia. W ten sposób optymalizowano poszczególne zespoły i podzespoły układu chłodzenia, tak aby spełniał on swoją rolę w zakładanych warunkach eksploatacji i ekologizację emisji źródeł energii (paliwa) i szkodliwych substancji.

Słowa kluczowe: *użytkowanie pojazdów ciężkich, zwalniacz pojazdu, wspomaganie hamowania, wymiana ciepła.*

1. Introduction

Modern trucks for the transport of cargo/humans, apart from the conventional braking system, are equipped with additional braking systems which can take the form of an electrodynamic brake, engine brake or hydraulic brake [5, 9, 10, 12]. When driving on long and steep road sections, the vehicle often uses the service brake, which can lead to increased heat load and serious wear to the brake system. This is a dangerous phenomenon affecting the decrease in the efficiency of the mechanical brake due to the thermal recession of the system [3, 4, 28]. Hydraulic retarders, commonly used in commercial vehicles, are auxiliary devices that can reduce vehicle speed by converting the vehicle's mechanical energy into heat energy absorbed by the retarder's working medium that is oil.

Compared to other auxiliary equipment, hydraulic retarders have many advantages such as low weight, high braking torque, long operating hours, good thermal diffusivity and zero environmental pollution. The design and principle of operation of hydraulic retarders/intarders in the literature is quite well described [14, 23, 24]. Nevertheless, the development of new technologies affects the efficiency

of these devices and their characteristics are being investigated by many researchers [18, 26, 35]. It is noted that they are mandatory in Europe for vehicles with a gross vehicle weight (GVW) of over 16 tones. Currently, they are also installed in vehicles with smaller GVW parameters [8]. These units are characteristic heat exchangers, the maintenance of which at the required level of efficiency also depends on the powertrain cooling system's capacity [17, 25, 29, 38]. Operation of the hydraulic retarder is periodic and its cooling system is connected to the cooling system of the traction engine.

The capacity of the cooling system must be selected in such a way as to meet the assumed operating conditions of the truck. When considering the thermal capacity of the mass of the powertrain and the efficiency of the cooling system, the thermal capacity received from the retarder over a period of 12 minutes of its operation shall be estimated. This time has been estimated on the basis of average downward slopes existing on the European roads [24]. This requires designers to develop efficient equipment.

Great emphasis was placed on modelling traction performance characteristics versus technical parameters and geometry of rotating elements of the retarder [6, 13, 30], as well as testing the medium flow field structure [23, 27]. Numerous studies have been devoted to mathematical modelling in the analyses of the effect of rotor speed and fill ratio on retarder output torque [15, 31, 32, 33, 34]. Mathematical models are based on complex differential equations of flows, which are often solved using numerical methods [12, 21, 36]. With the use of numerical models in computer simulations it is possible to make certain verifications of the flow rate or the average speed of the working fluid, as well as the braking torque - which was the subject of papers [1, 2, 20, 22]. In the analyzed literature [11, 19, 35, 37], the majority of problems concern the evaluation of retarders efficiency and their optimization, while the influence of the design heat output of the radiator on the braking efficiency of hydraulic retarder was rather simplified. Taking into account the design thermal capacity of the radiator in the coupled system of hydraulic retarder cooling system, main engine cooling system and engine oil cooling system, a detailed assessment of the thermal capacity of these units is required. Each of the mentioned devices is filled with a cooling medium with different thermo-physical properties. The entire system works correctly if none of the cooling media exceeds the permissible operating temperatures [24, 29].

The authors of this paper focused their attention on estimating the capacity of the main powertrain radiator, assuming that its efficiency can be reduced to about 15% due to local obstructions and contamination of the system. The aim of this paper is to select a main radiator working with multiple sub-systems and a hydraulic retarder. Achievement of the goal required the authors to develop a mathematical model, which was subject to simulation verification for two refrigerating media. The effects of ambient temperature, the volume of liquid in the retarder system, the size of the engine main radiator and the pump capacity in the cooling system on the temperature level of cooling media were analyzed. On the basis of the conducted tests and the obtained results, the direction of optimization of the adopted cooling system was indicated so that it complies with the requirements [7] under the assumed operating conditions. The obtained heat exchange coefficients were verified on a test stand, using the example of the 6C107[1] engine.

The presented calculations are only an example of how this model can be used for practical purposes, the versatility of which lies in the possibility of extending the powertrain cooling system with specific components.

2. Characteristics of the operating conditions of the vehicle under analysis

For the analytical tests, a propulsion system of a road vehicle powered by a diesel engine type 6C107, with a capacity of 6.53 dm^3 and power of 92.5 kW/2600 rpm, produced in Poland, was adopted [33]. It was assumed that the vehicle is equipped with a Voith R120-4 hydraulic retarder. The data of this retarder were used to estimate the coefficients characterizing the heat exchange in this device [27, 31]. The structure of the cooling system to be tested is illustrated in Fig. 1.

The engine cooling system was protected by a bellows thermostat with the characteristics shown in Fig. 2 [34].

During the simulation tests it was assumed that the vehicle was driving on a horizontal road, emitting energy into the cooling system in the amount of q_{dost} = 29.5 kW. After 50 minutes, the vehicle starts to descend on a gradient of i = 7%. The descent lasts 12 minutes. The vehicle then continues to run on flat ground. The initial time of

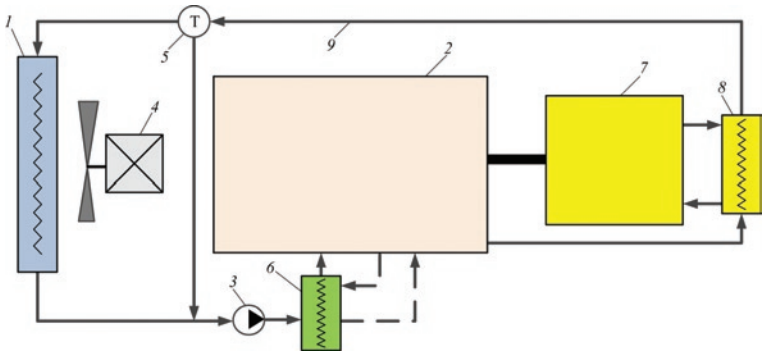


Fig. 1. Structure of the cooling system of a road vehicle powertrain equipped with a retarder: 1 – liquid-air cooler, 2 – engine, 3 – circulation pump of cooling liquid, 4 – fan of the cooler with electric drive, 5 – thermostat, 6 – engine oil cooler, 7 – retarder, 8 – retarder cooler, 9 – ducts

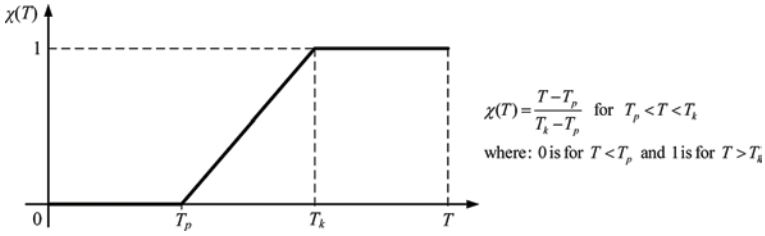


Fig. 2. Characteristics of the bellows thermostat for engine protection 6C107, where T_p – temperature at the beginning of thermostat opening, T_k – temperature of full thermostat opening

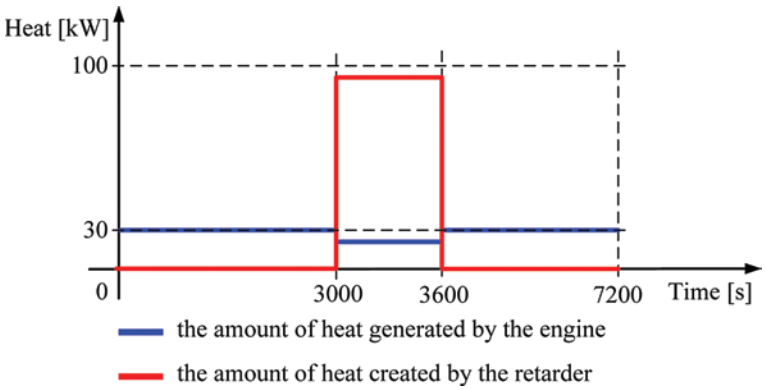


Fig. 3. Trend of thermal loads in the powertrain of the vehicle under analysis

1 The 6C107 engine is a design developed based on 400 Leyland

50 minutes of constant speed driving causes the temperatures in the engine cooling system to be close to asymptotic levels under the assumed operating conditions [36, 39]. The trend of thermal loads in the cooling system of the powertrain was assumed as shown in Fig. 3.

3. Powertrain heat exchange model

A diagram of the powertrain model is shown in Figure 4.

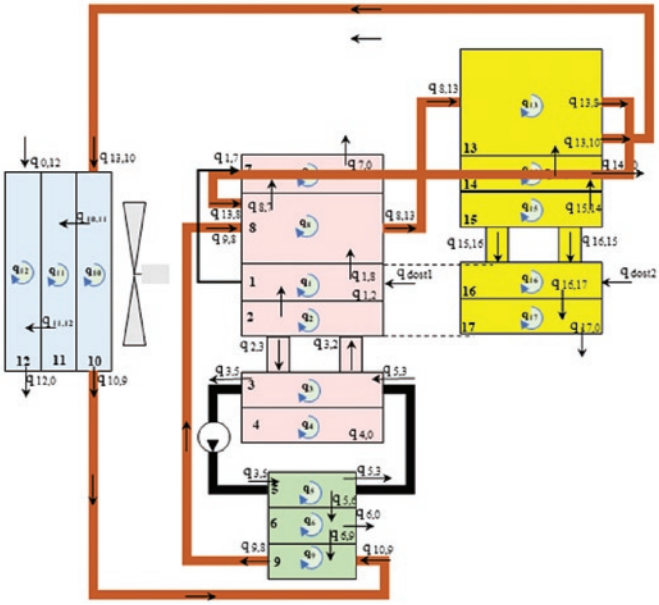


Fig. 4. Block diagram of engine with water-air cooler, engine oil cooler, retarder and thermostat; where: 0 – ambient environment, 1 – hot engine components, 2 – oil in the engine block, 3 – oil in the oil sump, 4 – oil sump, 5 – oil in the oil cooler, 6 – body of the oil cooler, 7 – engine body, 8 – water in the engine block, 9 – water in the oil cooler, 10 – water in the main cooler, 11 – body of the main cooler, 12 – water in the retarder cooler, 13 – water in the retarder cooler, 14 – body of the retarder cooler, 15 – retarder oil in the cooler, 16 – retarder oil in the retarder, 17 – retarder body

The thermal balance of the model is described by the following equations:

$$q_1 = q_{dost1} - q_{1,2} - q_{1,7} - q_{1,8} \tag{1}$$

$$q_2 = q_{1,2} + q_{3,2} - q_{2,3} \tag{2}$$

$$q_3 = q_{3,2} + q_{5,3} - q_{3,2} - q_{3,5} - q_{3,4} \tag{3}$$

$$q_4 = q_{3,4} - q_{4,0} \tag{4}$$

$$q_5 = q_{3,5} - q_{5,3} - q_{5,6} \tag{5}$$

$$q_6 = q_{5,6} - q_{6,0} - q_{6,9} \tag{6}$$

$$q_7 = q_{1,7} + q_{8,7} - q_{7,0} \tag{7}$$

$$q_8 = q_{1,8} + q_{9,8} + q_{13,8} - q_{8,7} - q_{8,13} \tag{8}$$

$$q_9 = q_{6,9} + q_{10,9} - q_{9,8} \tag{9}$$

$$q_{10} = q_{13,10} - q_{10,11} - q_{10,9} \tag{10}$$

$$q_{11} = q_{10,11} - q_{11,12} \tag{11}$$

$$q_{12} = q_{11,12} + q_{0,12} - q_{12,0} \tag{12}$$

$$q_{13} = q_{14,13} + q_{8,13} - q_{13,10} - q_{13,8} \tag{13}$$

$$q_{14} = q_{15,14} - q_{14,13} - q_{14,0} \tag{14}$$

$$q_{15} = q_{16,15} - q_{15,16} - q_{15,14} \tag{15}$$

$$q_{16} = q_{dost2} + q_{15,16} - q_{16,15} - q_{16,17} \tag{16}$$

$$q_{17} = q_{16,17} - q_{17,0} \tag{17}$$

Equations (1) to (17) in the differential notation take the following form:

$$c_1\dot{T}_1 = c_{dost1} - c_{1,2}(T_1 - T_2) - c_{1,7}(T_1 - T_7) - c_{1,8}(T_1 - T_8) \tag{18}$$

$$c_2\dot{T}_2 = c_{1,2}(T_1 - T_2) + c_{3,2}(T_3) - c_{2,3}(T_2) \tag{19}$$

$$c_3\dot{T}_3 = c_{2,3}(T_2) + c_{5,3}(T_5) - c_{3,2}(T_3) - c_{3,5}(T_3) - c_{3,4}(T_3 - T_4) \tag{20}$$

$$c_4\dot{T}_4 = c_{3,4}(T_3 - T_4) - c_{4,0}(T_4 - T_0) \tag{21}$$

$$c_5\dot{T}_5 = c_{3,5}(T_3) - c_{5,3}(T_3) - c_{5,6}(T_5 - T_6) \tag{22}$$

$$c_6\dot{T}_6 = c_{5,6}(T_5 - T_6) - c_{6,0}(T_6 - T_0) - c_{6,9}(T_6 - T_9) \tag{23}$$

$$c_7\dot{T}_7 = c_{8,7}(T_8 - T_7) + c_{1,7}(T_1 - T_7) - c_{7,0}(T_7 - T_0) \tag{24}$$

$$c_8\dot{T}_8 = c_{1,8}(T_1 - T_8) + \chi c_{9,8}(T_9) + (1-\chi)c_{13,8}(T_{13}) - c_{8,7}(T_8 - T_7) - c_{8,13}(T_8) \tag{25}$$

$$c_9\dot{T}_9 = c_{6,9}(T_6 - T_9) + \chi c_{10,9}(T_{10}) - \chi c_{9,8}(T_9) \tag{26}$$

$$c_{10}\dot{T}_{10} = \chi c_{13,10}(T_{13}) - \chi c_{10,9}(T_{10}) - c_{10,11}(T_{10} - T_{11}) \tag{27}$$

$$c_{11}\dot{T}_{11} = c_{10,11}(T_{10} - T_{11}) - c_{11,12}(T_{11} - T_{12}) \tag{28}$$

$$c_{12}\dot{T}_{12} = c_{11,12}(T_{11} - T_{12}) + c_{0,12}(T_0) - c_{12,0}(T_{12}) \tag{29}$$

$$c_{13}\dot{T}_{13} = c_{14,13}(T_{14} - T_{13}) + c_{8,13}(T_8) - \chi c_{13,10}(T_{13}) - (1-\chi)c_{13,8}(T_{13}) \tag{30}$$

$$c_{14}\dot{T}_{14} = c_{15,14}(T_{15} - T_{14}) - c_{14,13}(T_{14} - T_{13}) - c_{14,0}(T_{14} - T_0) \tag{31}$$

$$c_{15}\dot{T}_{15} = c_{16,15}(T_{16}) - c_{15,16}(T_{15}) - c_{15,14}(T_{15} - T_{14}) \tag{32}$$

$$c_{16}\dot{T}_{16} = c_{dost2} + c_{15,16}(T_{15}) - c_{16,15}(T_{16}) - c_{16,17}(T_{16} - T_{17}) \tag{33}$$

$$c_{17}\dot{T}_{17} = c_{16,17}(T_{16} - T_{17}) - c_{17,0}(T_{17} - T_0) \tag{34}$$

The values of c_i and c_{ij} factors were estimated on the basis of technical documentation of individual elements of the powertrain. The c_i

values describe the unit thermal capacities of the distinguished elements and equal the product of the mass of the element [kg] and its specific heat [kJ/kg·K]. The c_{ij} values describe the heat transfer conditions on the surface of the mentioned element and are equal to the product of the heat transfer surface of the element [m^2] and the heat transfer coefficient on the surface [kJ/m^2·K] (see Table 1).

The model of heating the engine itself without a retarder has been verified experimentally on the engine dynamometer in the Laboratory of Combustion Engines of the Wroclaw University of Technology and Science (Fig. 5). The results obtained from the measurement of the selected parameters describing the operation of the cooling system were used for determining the threshold conditions in the simulation model. Matlab-Simulink package was used for simulation tests (Fig. 6).

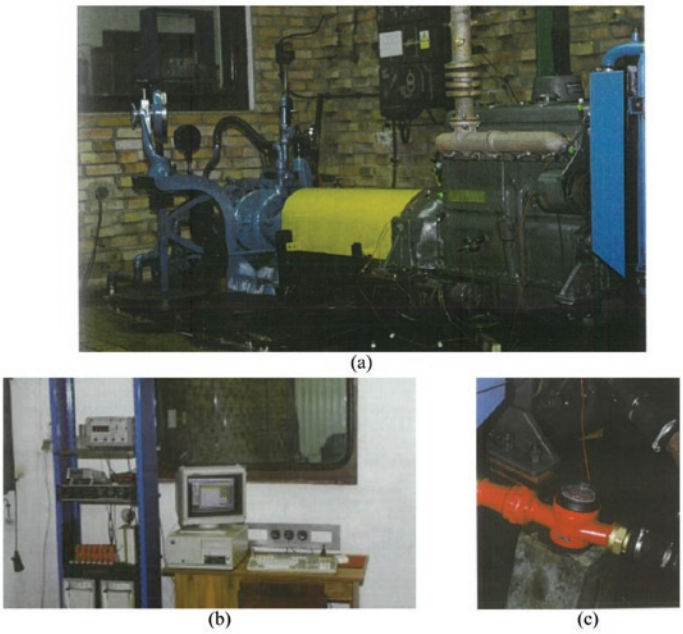


Fig. 5. Pictures of the test stand: (a) view of the 6C107 engine on the engine dynamometer of the Wroclaw University of Technology and Science, (b) computerized test bench on a dynamometer, with a temperature measurement module, (c) view of the engine block coolant flow meter

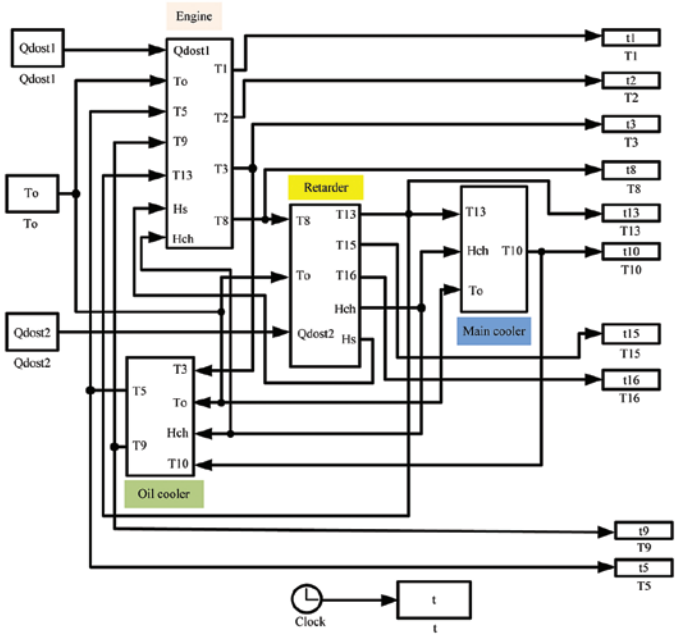


Fig. 6. Matlab-Simulink model

4. Results of simulation tests

Simulation tests were preceded by the development of a test plan. All components of the powertrain, i.e. temperatures from T_1 to T_{16}, distinguished in the temperature model, according to the legend to Fig. 6. A series of tests was carried out taking into account different ambient temperatures, heat exchange surfaces of the radiator, increasing the thermal capacity of the system. The tests used a cooling liquid based on:

a) water-filled engine cooling system,
b) glycol fluid-filled cooling system.

The continuous line indicates temperature fluctuations for the analysed subsystems – T_3, T_8 and T_{16}. The dotted line indicates threshold values in the analysed subsystems.

4.1. Simulation tests of a cooling system at different ambient temperatures

The cooling system tests were carried out at ambient temperatures (T_{ot}) in the range of 253K to 313K every 10K steps. Examples of results are shown in Fig. 7 and Fig. 8. According to them, the system can operate safely in the ambient temperature range up to T_{ot} = 306K (33ºC) (the highest ambient temperatures in Poland reach 313K (40ºC). The improvement of the situation can be achieved by enlarging e.g. the cooler of the engine block coolant.

4.2. Simulation tests of the cooling system for the case of an increase in the heat exchange area of the main cooler

The tests were carried out for three variants of cooler sizes: increased by 10%, by 30% and by 50%, at ambient temperature T_{ot} = 313K (40ºC). The results are presented in Fig. 9.

As can be seen from the presented results, in a glycol-filled system, the use of a cooler with capacity increased by 50% effectively protects all media against exceeding the permissible temperatures, especially the temperature of oil in the retarder. The use of a 50% larger main cooler led to a reduction in oil temperature of about 12K. However, this is not an optimal solution. It is worth investigating the impact that increasing the capacity of the pump in the main cooling system can have on overall capacity.

4.3. Simulation tests of the vehicle cooling system for the case of the cooler increased by 50% and increased coolant flow

The studies were carried out, as in previous cases, for water and glycol at increased flows by 50% and by 100% at ambient temperature T_{ot} = 313K (40ºC). The results are shown in Fig. 10.

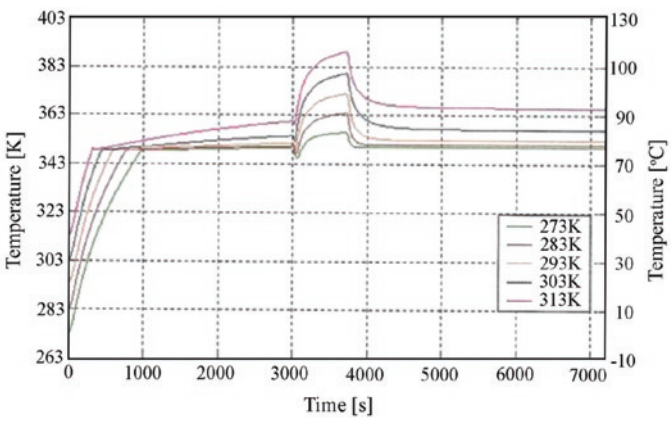


Fig. 7. Trend of cooling liquid temperature changes in engine block at different ambient temperatures (experiment)

Table 1. List of coefficients describing the engine thermal state

No.	A	B	C	D
1	c_i [J/K]	$c_{i,j}$ [W/K]	$c_{i,0}$ [J/K]	other parameters
2	$c_1 = 87990$	$c_{1,2} = 16$	$c_{4,0} = 1.2$	$T_0 = 293.15$ [K]
3	$c_2 = 11750$	$c_{1,7} = 200$	$c_{6,0} = 10$	$q_{dost1} = 29500$ [W]
4	$c_3 = 1796$	$c_{1,8} = 998$	$c_{7,0} = 60$	$q_{dost2} = 91263$ [W]
5	$c_4 = 5560$	$c_{2,3} = 20$	$c_{12,0} = 13000$	$T_p = 348$ [K]
6	$c_5 = 6080$	$c_{3,4} = 097$	$c_{14,0} = 12.5$	$T_k = 358$ [K]
7	$c_6 = 1800$	$c_{3,5} = 190$	$c_{17,0} = 50$	-
8	$c_7 = 448148$	$c_{5,6} = 72$	-	-
9	$c_8 = 20000$	$c_{6,9} = 100$	-	-
10	$c_9 = 20950$	$c_{8,7} = 236$	-	-
11	$c_{10} = 41900$	$c_{9,8} = 5866$	-	-
12	$c_{11} = 19250$	$c_{9,10} = 5866$	-	-
13	$c_{12} = 91$	$c_{10,11} = 955.6$	-	-
14	$c_{13} = 20960$	$c_{11,12} = 1240$	-	-
15	$c_{14} = 4500$	$c_{13,18} = 5866$	-	-
16	$c_{15} = 6840$	$c_{13,10} = 5866$	-	-
17	$c_{16} = 3040$	$c_{14,13} = 583$	-	-
18	$c_{17} = 24750$	$c_{15,14} = 456$	-	-
19	-	$c_{16,15} = 5866$	-	-
20	-	$c_{16,17} = 847$	-	-

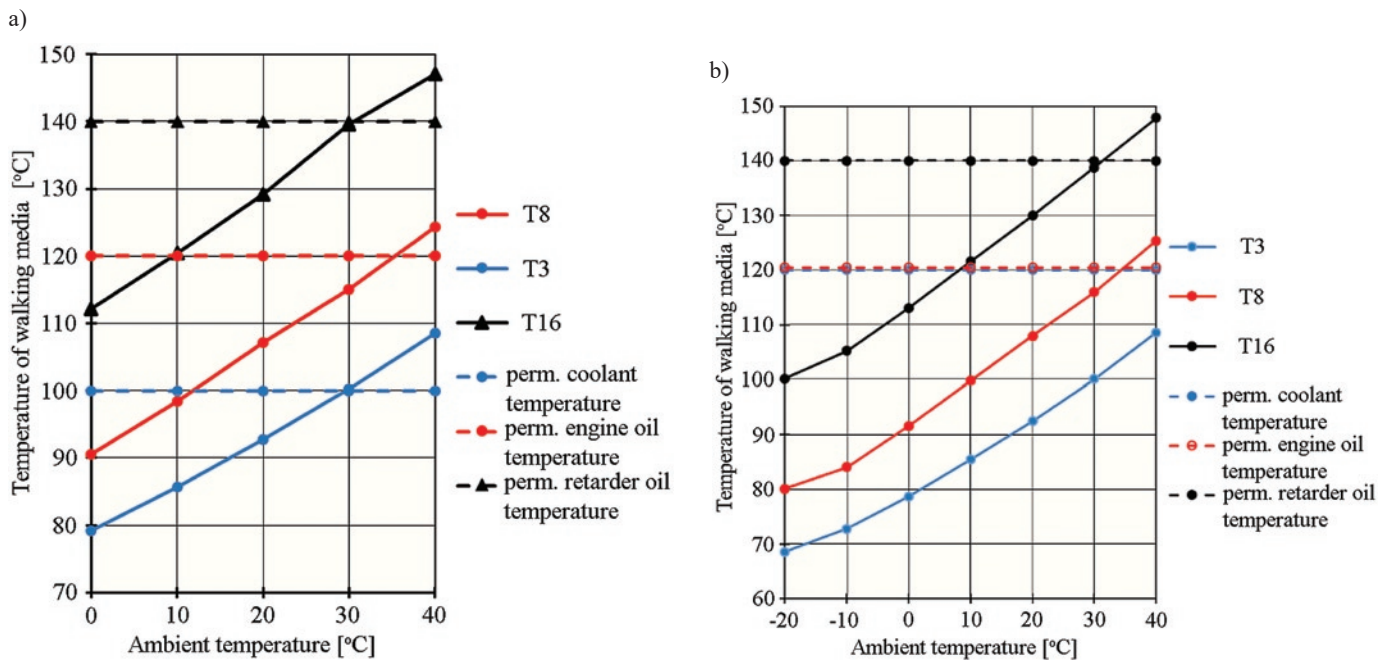


Fig. 8. Temperature of cooling media T_8, T_3, T_{16} versus ambient temperature T_{ot} (experiment): (a) for water, (b) for glycol

Research results show that an effective solution to the problem of temperature reduction in the retarder cooling system can be to increase the pump capacity by 50% or even only by 25% [15, 16].

4.4. Simulation tests of the vehicle's cooling system when an additional coolant tank is used to increase the thermal capacity of the system

The tests were carried out for tanks with a capacity of 20 and 40 dm^3 at an ambient temperature of T_{ot} = 313K (40ºC). The tank was located in the so-called large cooling circuit between the retarder cooler and the main engine cooler. It was assumed that the cooling system is filled with water. The results are shown in Fig. 11.

Equipping the cooling system with an additional coolant tank with a capacity of up to 40 dm^3 causes a slight decrease in system temperature. This solution is ineffective.

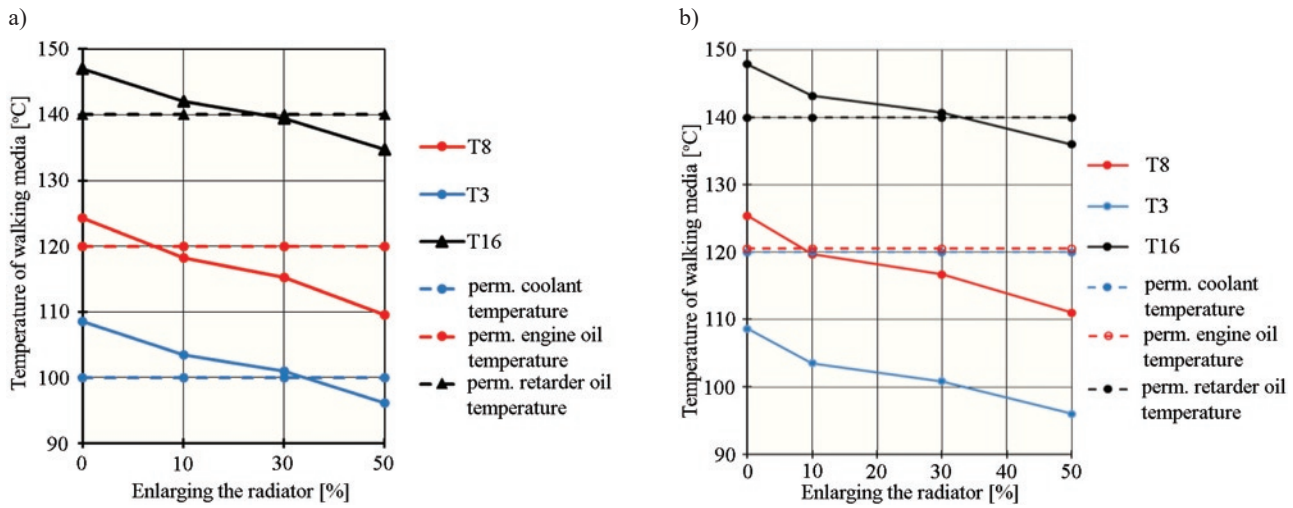


Fig. 9. Temperature of cooling media T_8, T_3, T_{16} versus the size of the radiator T_{ot} (simulation): (a) for water, (b) for glycol

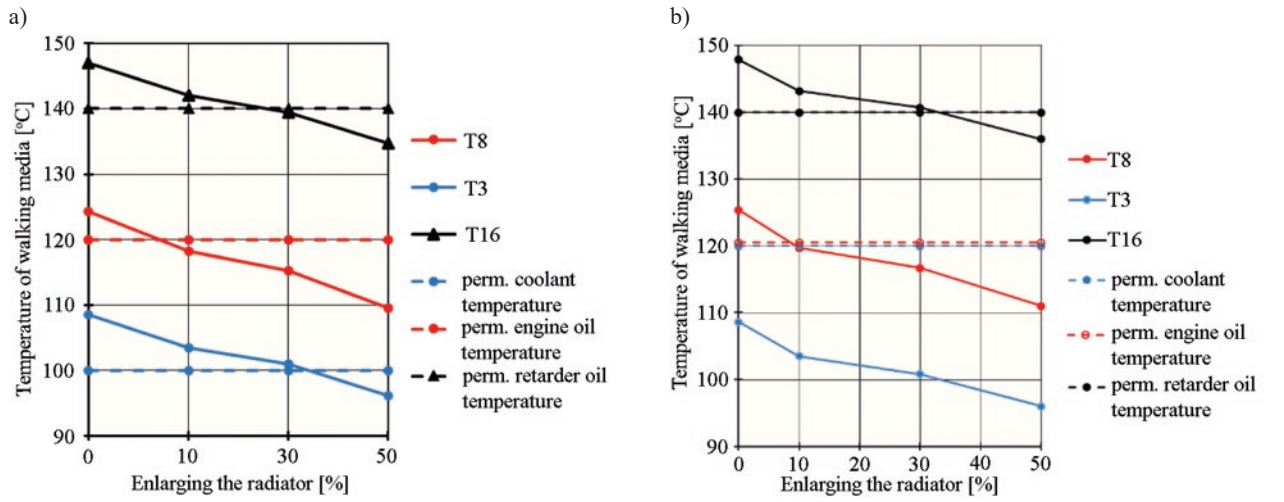


Fig. 10. Temperature of cooling media T_8, T_3, T_{16} versus coolant flow T_{ot} (simulation): (a) for water, (b) for glycol

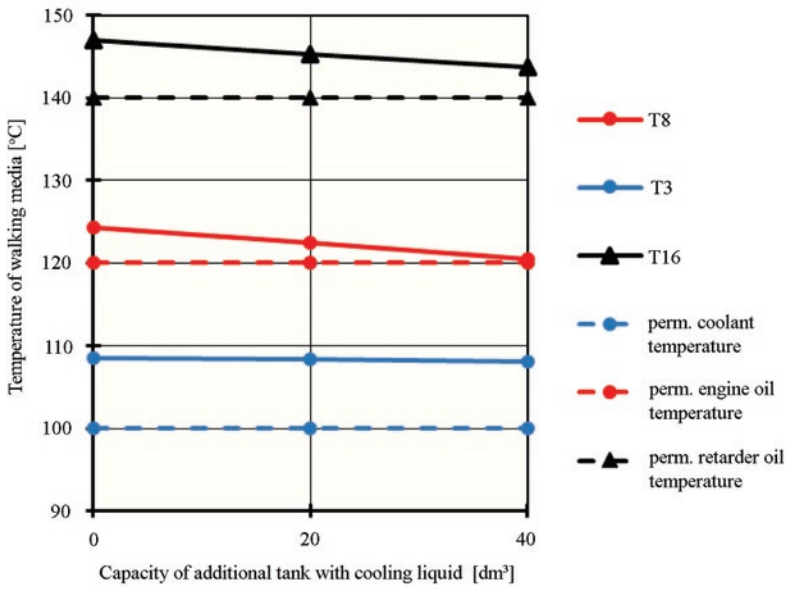


Fig. 11. The dependence of cooling media temperatures on the size of the additional coolant tank (simulation)

5. Conclusion

The aim of the study was to analyze the dynamics of heat exchange in a vehicle equipped with a hydraulic retarder. This device emits large energy flows into the cooling system of the braked vehicle. Therefore, the cooling system should take into account the retarder operation. The requirements in this respect are defined by the relevant EC regulations. On the basis of the analysis of the cooling system structure, a structural and computational model was built. The model of the cooling system is based on a set of 17 balance equations which, according to Newton's principle, were transformed into a set of 17 differential equations describing temperature changes of the distinguished elements versus time. The Matlab-Simulink package was used to solve the equation system. In the study, analyses of the influence of working conditions and constructional conditions on the behaviour of the system in the assumed scenario of its operation were carried out. The results of the work allowed to indicate the direction in which the design of the cooling system of the drive unit should be upgraded so that it could perform its function even in the most difficult working conditions.

References

1. Abe K, Kondoh T, Fukumura K, Kojima M. Three-dimensional simulation of the flow in a torque converter. SAE 1991, https://doi.org/10.4271/910800.
2. Albertz D, Dappen S, Henneberger G. Calculation of the 3D nonlinear eddy current field in moving conductors and its application to braking systems. IEEE Transactions on Magnetics 1996; 32(3): 768-771, https://doi.org/10.1109/20.497353.
3. Baranowski P, Damaziak K, Malachowski J. Brake system studies using numerical methods. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(4): 337-341.
4. Baranowski P, Damaziak K, Malachowski J, Mazurkiewicz L, Kastek M, Polakowski H, Piatkowski T. Experimental and numerical tests of thermomechanical processes occurring on brake pad lining surface. Surface Effects and Contact Mechanics 2011; 10: 15-24, https://doi.org/10.2495/SECM110021.
5. Christoffersen S, Wallingford J, Greenlees B. Heavy truck engine retarders. Testing and theory. SAE 2011, https://doi.org/10.4271/2011-01-0280.
6. Dong Y, Korivi V, Attibele P, Yuan Y. Torque converter CFD engineering part I: Torque ratio and K factor improvement through stator modifications. SAE 2002, https://doi.org/10.4271/2002-01-0883.
7. ECE Regulation No. 13, Uniform provisions concerning the approval of vehicles of categories M, N and O with regard to braking.
8. Gohring E, Glasner EC, Povel R. Engine braking systems and retarders - an overview from European standpoint. SAE 1992, https://doi.org/10.4271/922451.
9. Habib G. The present status of electro-magnetic retarders in commercial vehicles. SAE 1992, https://doi.org/10.4271/922450.
10. Haiss G. Demand criteria on retarders. SAE 1992; https://doi.org/10.4271/922453.
11. Heisler H. Advanced Vehicle Technology. Second Edition. Butterworth-Heinemann: Elsevier Ltd., 2002.
12. Jeyakumar S, Sasikumar M. Computational fluid dynamics simulation of hydraulic torque converter for performance characteristics prediction. International Journal of Scientific Research in Science, Engineering and Technology 2017; 3(6): 402-408.
13. Jia Y H. Dynamic simulation research of hydrodynamic retarder in brake process. Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE) 2011: 1193-1196.
14. Kazmierczak A, Krakowian K, Wrobel. Doppler Laser Vibrometry in combustion engine's diagnosis. Przeglad Elektroniczny 2010; 86(10): 147-149.
15. Kern J, Ambros P. Concepts for a controlled optimized vehicle engine cooling system. SAE 1997, https://doi.org/10.4271/971816.
16. Kesy A, Kadziela A. Construction optimization of hydrodynamic torque converter with application of genetic algorithm. Archives of Civil and Mechanical Engineering 2011; 11(4): 905-920, https://doi.org/10.1016/S1644-9665(12)60086-7.
17. Kwasniowski S, Sroka Z. Modeling dynamics of heat exchange in cooling, heating and air conditioning systems of vehicles and working machines. Wroclaw: Series report SPR No. 075/97, Wroclaw University of Technology, 1998.
18. Lei Y, Song P, Zheng H, Fu Y, Li X, Song B. Application of fuzzy logic in constant speed control of hydraulic retarder. Advances in Mechanical Engineering 2017; 9(2): 1-11, https://doi.org/10.1177/1687814017690956.
19. Li J, Tan G, Ji Y, Zhou Y, Liu Z, Xu Y. Design and simulation analysis for an integrated energy-recuperation retarder. SAE Technical Paper 2016, https://doi.org/10.4271/2016-01-0458.
20. Li R, Yang J, Zhang W. Simulation Study of the Vehicle Hydraulic Retarder, International Journal of Control and Automation 2015; 8(2): 263-280, https://doi.org/10.14257/ijca.2015.8.2.26.
21. Liu C Y, Jiang K J, Zhang Y. Design and use of an eddy current retarder in an automobile. International Journal of Automotive Technology 2011; 12(4): 611-616, https://doi.org/10.1007/s12239-011-0071-3.
22. Marechal Y, Meunier G. Computation of 2D and 3D eddy currents in moving conductors of electromagnetic retarders. IEEE Transactions on Magnetics1990; 26(5): 2382-2384, https://doi.org/10.1109/20.104738.
23. Mu H, Yan Q, Wei W. Study on influence of inlet and outlet flow rates on oil pressures and braking torque in a hydrodynamic retarder. International Journal of Numerical Methods for Heat & Fluid Flow 2017; 27(11): 2544-2564, https://doi.org/10.1108/HFF-10-2016-0428.
24. Pandey S N, Khaliq A, Zaka M Z, Saleem M S, Afzal M. Retarder used as braking system in heavy vehicles - a review. International Journal Mechanical Engineering and Robotic Research 2015; 4(2): 86-90.
25. Peng Z. A Study into the technology of development of hydraulic. Xi'an: Chang'an University, 2008.
26. Pernestål A, Nyberg M, Warnquist, H. Modeling and inference for troubleshooting with interventions applied to a heavy truck auxiliary braking system. Engineering Applications of Artificial Intelligence 2012; 25(4): 705-719, https://doi.org/10.1016/j.engappai.2011.02.018.
27. Song B, Lv J, Liu Y, Kong F. The simulation and analysis on engine and hydraulic retarder continual braking performance of the tracked vehicle on long downhill. Proceedings 9th International Conference on Electronic Measurement & Instruments 2009: 3-928-3-931, https://doi.org/10.1109/ICEMI.2009.5274169.
28. Sarkar S, Rathod P P. Review paper on thermal analysis of ventilated disc brake by varying design parameters. International Journal of Engineering Research & Technology 2013; 2(12): 1077-1081.
29. Tan G, Guo X, Yang T. Simulation based heavy truck driveline components thermal analysis system. The 9th International Conference on, Electronic Measurement & Instruments, ICEMI'2009: 4-675-4-680, https://doi.org/10.1109/ICEMI.2009.5274676.
30. Tan G, Guo X. The modeling and performance analysis of the retarder thermal management system. SAE 2012, https://doi.org/10.4271/2012-01-1929.
31. Wambsganss M W. Thermal management in heavy vehicles: a review identifying issues and research requirements. Argonne National Lab., IL (US), No. ANL/ET/CP-98208, 1999.
32. Wangand G, Shan S. Review of meta modeling techniques in support of engineering design optimization. Journal of Mechanical Design, Transactions of the ASME 2007; 129(4): 370-380, https://doi.org/10.1115/1.2429697.
33. Wrzecioniarz P, Kwasniowski S, Jamroziak K. Criterion for choosing the power unit cooling system of evacuation tractor. Czasopismo Techniczne Mechanika 1998; 95(5-M): 83-93.
34. Wrzecioniarz P, Kwasniowski S, Jamroziak K. The concept of the cooling system for the road tractor unit. Pojazdy samochodowe: problemy

rozwoju jakości, VI Międzynarodowa Konferencja Naukowo-Techniczna "Autoprogres'98" 1998; T.2: 127-134.
35. Xin Q. Diesel engine system design. New Delhi: Woodhead Publishing, 2011, https://doi.org/10.1533/9780857090836.
36. Yang J, Yi F, Wang J. Model-based adaptive control of eddy current retarder. Proceeding of the 30th Chinese Control And Decision Conference (2018 CCDC) 2018; 1889-1891, https://doi.org/10.1109/CCDC.2018.8407434.
37. Zheng H, Lei Y, Song P. Design of a filling ratio observer for ahydraulic retarder: An analysis ofvehicle thermal management and dynamic braking system. Advances in Mechanical Engineering 2016; 8(10): 1-8, https://doi.org/10.1177/1687814016674098.
38. Zheng H, Lei Y, Song P. Hydraulic retarders for heavy vehicles: Analysis of fluid mechanics and computational fluid dynamics on braking torque and temperature rise. International Journal of Automotive Technology 2017; 18(3): 387-396, https://doi.org/10.1007/s12239-017-0039-z.
39. Zhou L, Tan G, Guo X, Chen M, Ji K, Li Z. Yang Z. Study of energy recovery system based on organic rankine cycle for hydraulic retarder. SAE 2016, https://doi.org/10.4271/2016-01-0239.

Krzysztof JAMROZIAK
Stanislaw KWASNIOWSKI
Mariusz KOSOBUDZKI
Pawel ZAJAC
Faculty of Mechanical Engineering,
Wroclaw University of Science and Technology
ul. Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland

E-mails: krzysztof.jamroziak@pwr.edu.pl, stanislaw.kwasniowski@pwr.edu.pl, mariusz.kosobudzki@pwr.edu.pl, pawel.zajac@pwr.edu.pl

Article citation info:
JAMROZIAK K, KWASNIOWSKI S, KOSOBUDZKI M, ZAJAC P. Analysis of heat exchange in the powertrain of a road vehicle with a retarder. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 577–584, http://dx.doi.org/10.17531/ein.2019.4.6.

Krzysztof JAMROZIAK
Stanislaw KWASNIOWSKI
Mariusz KOSOBUDZKI
Pawel ZAJAC

ANALYSIS OF HEAT EXCHANGE IN THE POWERTRAIN OF A ROAD VEHICLE WITH A RETARDER

ANALIZA WYMIANY CIEPŁA W UKŁADZIE NAPĘDOWYM POJAZDU DROGOWEGO Z RETARDEREM

The paper presents a heat exchange model for the cooling system of any complex, physical system. Verification of the correctness of the theoretical model was carried out on the example of a vehicle with a combustion engine and additionally equipped with a hydraulic retarder. The results of laboratory tests, which were carried out on an engine test bench, were also performed for the above mentioned powertrain, so as to compare the results of modelling with the results of the tests. Determining the operating parameters of the components of the cooling system aimed at protecting the entire powertrain against overheating is a key task. Theoretical analysis of heat exchange in the powertrain of a road vehicle was carried out, with particular emphasis on the hydraulic retarder (a device braking the vehicle during a descent on roads with a high gradient of the road, mandatory according to the ADR convention). The subject of the study was a mathematical model of a complex cooling system developed by the authors, described by means of balance equations and differential equations. This model was tested with the use of the Matlab-Simulink suite for given load parameters of the cooling system, which were used in tests on an engine test bench. The values of coefficients describing the thermal state of the powertrain were obtained. Simulations were performed for different variants of technical parameters of the expanded cooling system. In this way, individual units and components of the cooling system were optimized so that it fulfilled its role in the assumed operating conditions and the ecologization of emission of energy sources (fuel) and harmful substances.

Keywords: *heavy-duty vehicle operation, vehicle retarder, brake support, heat exchange.*

Praca zawiera model wymiany ciepła układu chłodzenia dla złożonego, dowolnego układu fizycznego. Weryfikację poprawności modelu teoretycznego przeprowadzono na przykładzie pojazdu z silnikiem spalinowym oraz dodatkowo wyposażonym w retarder hydrauliczny. Wyniki badań laboratoryjnych, które przeprowadzono na hamowni również wykonano dla w/w zespołu napędowego, tak aby porównać wyniki modelowania z wynikami badań. Ustalenie parametrów eksploatacji elementów układu chodzenia, którego celem jest zabezpieczenie całego układu napędowego przed przegrzaniem to kluczowe zadanie. Przeprowadzono teoretyczną analizę wymiany ciepła w układzie napędowym pojazdu drogowego ze szczególnym uwzględnieniem zwalniacza hydraulicznego (urządzenia hamującego pojazd podczas zjazdu na drogach o dużym pochyleniu jezdni, obowiązkowe wg konwencji ADR). Przedmiotem badań był opracowany przez autorów model matematyczny rozbudowanego układu chłodzenia opisany za pomocą równań bilansowych i równań różniczkowych. Model ten testowano z wykorzystaniem pakietu Matlab-Simulink dla zadanych parametrów obciążenia układu chłodzenia, które wykorzystywano w badaniach w stanowiskowych na hamowni silnikowej. Uzyskano wartości współczynników opisujących stan cieplny jednostki napędowej. Symulacje wykonano dla różnych wariantów parametrów technicznych rozbudowanego układu chłodzenia. W ten sposób optymalizowano poszczególne zespoły i podzespoły układu chłodzenia, tak aby spełniał on swoją rolę w zakładanych warunkach eksploatacji i ekologizację emisji źródeł energii (paliwa) i szkodliwych substancji.

Słowa kluczowe: *użytkowanie pojazdów ciężkich, zwalniacz pojazdu, wspomaganie hamowania, wymiana ciepła.*

1. Introduction

Modern trucks for the transport of cargo/humans, apart from the conventional braking system, are equipped with additional braking systems which can take the form of an electrodynamic brake, engine brake or hydraulic brake [5, 9, 10, 12]. When driving on long and steep road sections, the vehicle often uses the service brake, which can lead to increased heat load and serious wear to the brake system. This is a dangerous phenomenon affecting the decrease in the efficiency of the mechanical brake due to the thermal recession of the system [3, 4, 28]. Hydraulic retarders, commonly used in commercial vehicles, are auxiliary devices that can reduce vehicle speed by converting the vehicle's mechanical energy into heat energy absorbed by the retarder's working medium that is oil.

Compared to other auxiliary equipment, hydraulic retarders have many advantages such as low weight, high braking torque, long operating hours, good thermal diffusivity and zero environmental pollution. The design and principle of operation of hydraulic retarders/intarders in the literature is quite well described [14, 23, 24]. Nevertheless, the development of new technologies affects the efficiency

of these devices and their characteristics are being investigated by many researchers [18, 26, 35]. It is noted that they are mandatory in Europe for vehicles with a gross vehicle weight (GVW) of over 16 tones. Currently, they are also installed in vehicles with smaller GVW parameters [8]. These units are characteristic heat exchangers, the maintenance of which at the required level of efficiency also depends on the powertrain cooling system's capacity [17, 25, 29, 38]. Operation of the hydraulic retarder is periodic and its cooling system is connected to the cooling system of the traction engine.

The capacity of the cooling system must be selected in such a way as to meet the assumed operating conditions of the truck. When considering the thermal capacity of the mass of the powertrain and the efficiency of the cooling system, the thermal capacity received from the retarder over a period of 12 minutes of its operation shall be estimated. This time has been estimated on the basis of average downward slopes existing on the European roads [24]. This requires designers to develop efficient equipment.

Great emphasis was placed on modelling traction performance characteristics versus technical parameters and geometry of rotating elements of the retarder [6, 13, 30], as well as testing the medium flow field structure [23, 27]. Numerous studies have been devoted to mathematical modelling in the analyses of the effect of rotor speed and fill ratio on retarder output torque [15, 31, 32, 33, 34]. Mathematical models are based on complex differential equations of flows, which are often solved using numerical methods [12, 21, 36]. With the use of numerical models in computer simulations it is possible to make certain verifications of the flow rate or the average speed of the working fluid, as well as the braking torque - which was the subject of papers [1, 2, 20, 22]. In the analyzed literature [11, 19, 35, 37], the majority of problems concern the evaluation of retarders efficiency and their optimization, while the influence of the design heat output of the radiator on the braking efficiency of hydraulic retarder was rather simplified. Taking into account the design thermal capacity of the radiator in the coupled system of hydraulic retarder cooling system, main engine cooling system and engine oil cooling system, a detailed assessment of the thermal capacity of these units is required. Each of the mentioned devices is filled with a cooling medium with different thermo-physical properties. The entire system works correctly if none of the cooling media exceeds the permissible operating temperatures [24, 29].

The authors of this paper focused their attention on estimating the capacity of the main powertrain radiator, assuming that its efficiency can be reduced to about 15% due to local obstructions and contamination of the system. The aim of this paper is to select a main radiator working with multiple sub-systems and a hydraulic retarder. Achievement of the goal required the authors to develop a mathematical model, which was subject to simulation verification for two refrigerating media. The effects of ambient temperature, the volume of liquid in the retarder system, the size of the engine main radiator and the pump capacity in the cooling system on the temperature level of cooling media were analyzed. On the basis of the conducted tests and the obtained results, the direction of optimization of the adopted cooling system was indicated so that it complies with the requirements [7] under the assumed operating conditions. The obtained heat exchange coefficients were verified on a test stand, using the example of the 6C107[1] engine.

The presented calculations are only an example of how this model can be used for practical purposes, the versatility of which lies in the possibility of extending the powertrain cooling system with specific components.

2. Characteristics of the operating conditions of the vehicle under analysis

For the analytical tests, a propulsion system of a road vehicle powered by a diesel engine type 6C107, with a capacity of 6.53 dm^3 and power of 92.5 kW/2600 rpm, produced in Poland, was adopted [33]. It was assumed that the vehicle is equipped with a Voith R120-4 hydraulic retarder. The data of this retarder were used to estimate the coefficients characterizing the heat exchange in this device [27, 31]. The structure of the cooling system to be tested is illustrated in Fig. 1.

The engine cooling system was protected by a bellows thermostat with the characteristics shown in Fig. 2 [34].

During the simulation tests it was assumed that the vehicle was driving on a horizontal road, emitting energy into the cooling system in the amount of q_{dost} = 29.5 kW. After 50 minutes, the vehicle starts to descend on a gradient of i = 7%. The descent lasts 12 minutes. The vehicle then continues to run on flat ground. The initial time of

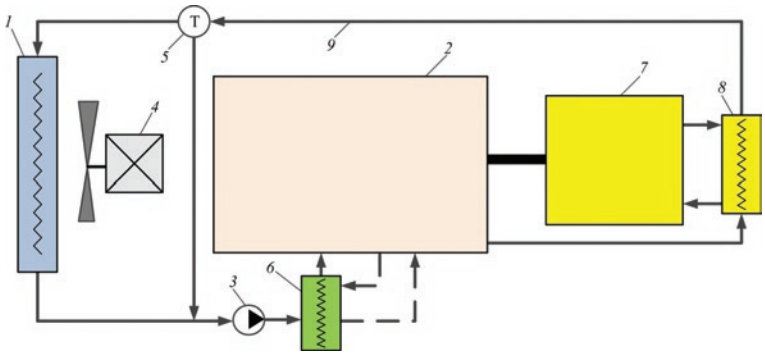


Fig. 1. Structure of the cooling system of a road vehicle powertrain equipped with a retarder: 1 – liquid-air cooler, 2 – engine, 3 – circulation pump of cooling liquid, 4 – fan of the cooler with electric drive, 5 – thermostat, 6 – engine oil cooler, 7 – retarder, 8 – retarder cooler, 9 – ducts

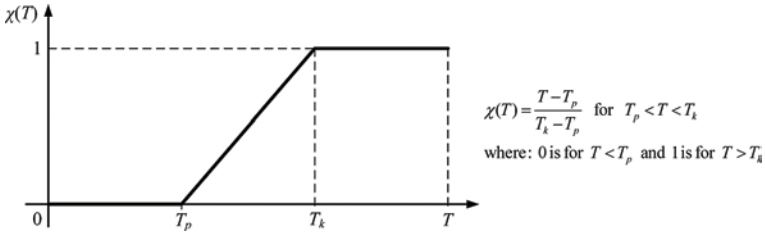


Fig. 2. Characteristics of the bellows thermostat for engine protection 6C107, where T_p – temperature at the beginning of thermostat opening, T_k – temperature of full thermostat opening

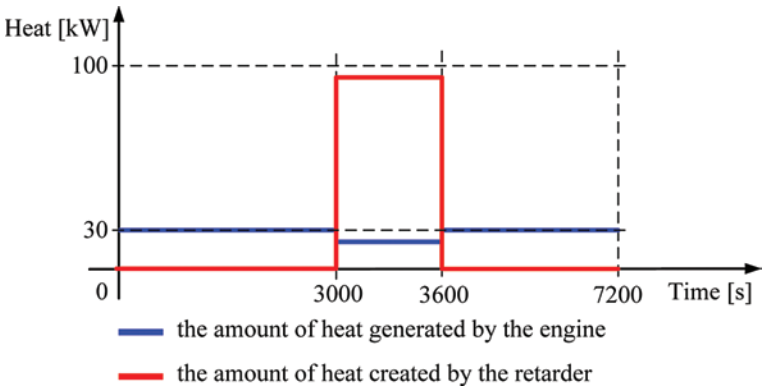


Fig. 3. Trend of thermal loads in the powertrain of the vehicle under analysis

1 The 6C107 engine is a design developed based on 400 Leyland

50 minutes of constant speed driving causes the temperatures in the engine cooling system to be close to asymptotic levels under the assumed operating conditions [36, 39]. The trend of thermal loads in the cooling system of the powertrain was assumed as shown in Fig. 3.

3. Powertrain heat exchange model

A diagram of the powertrain model is shown in Figure 4.

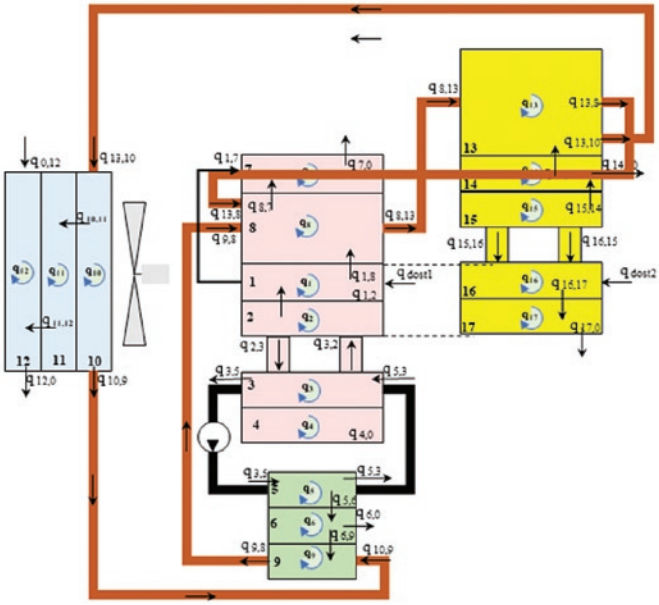


Fig. 4. Block diagram of engine with water-air cooler, engine oil cooler, retarder and thermostat; where: 0 – ambient environment, 1 – hot engine components, 2 – oil in the engine block, 3 – oil in the oil sump, 4 – oil sump, 5 – oil in the oil cooler, 6 – body of the oil cooler, 7 – engine body, 8 – water in the engine block, 9 – water in the oil cooler, 10 – water in the main cooler, 11 – body of the main cooler, 12 – water in the retarder cooler, 13 – water in the retarder cooler, 14 – body of the retarder cooler, 15 – retarder oil in the cooler, 16 – retarder oil in the retarder, 17 – retarder body

The thermal balance of the model is described by the following equations:

$$q_1 = q_{dost1} - q_{1,2} - q_{1,7} - q_{1,8} \tag{1}$$

$$q_2 = q_{1,2} + q_{3,2} - q_{2,3} \tag{2}$$

$$q_3 = q_{3,2} + q_{5,3} - q_{3,2} - q_{3,5} - q_{3,4} \tag{3}$$

$$q_4 = q_{3,4} - q_{4,0} \tag{4}$$

$$q_5 = q_{3,5} - q_{5,3} - q_{5,6} \tag{5}$$

$$q_6 = q_{5,6} - q_{6,0} - q_{6,9} \tag{6}$$

$$q_7 = q_{1,7} + q_{8,7} - q_{7,0} \tag{7}$$

$$q_8 = q_{1,8} + q_{9,8} + q_{13,8} - q_{8,7} - q_{8,13} \tag{8}$$

$$q_9 = q_{6,9} + q_{10,9} - q_{9,8} \tag{9}$$

$$q_{10} = q_{13,10} - q_{10,11} - q_{10,9} \tag{10}$$

$$q_{11} = q_{10,11} - q_{11,12} \tag{11}$$

$$q_{12} = q_{11,12} + q_{0,12} - q_{12,0} \tag{12}$$

$$q_{13} = q_{14,13} + q_{8,13} - q_{13,10} - q_{13,8} \tag{13}$$

$$q_{14} = q_{15,14} - q_{14,13} - q_{14,0} \tag{14}$$

$$q_{15} = q_{16,15} - q_{15,16} - q_{15,14} \tag{15}$$

$$q_{16} = q_{dost2} + q_{15,16} - q_{16,15} - q_{16,17} \tag{16}$$

$$q_{17} = q_{16,17} - q_{17,0} \tag{17}$$

Equations (1) to (17) in the differential notation take the following form:

$$c_1\dot{T}_1 = c_{dost1} - c_{1,2}(T_1 - T_2) - c_{1,7}(T_1 - T_7) - c_{1,8}(T_1 - T_8) \tag{18}$$

$$c_2\dot{T}_2 = c_{1,2}(T_1 - T_2) + c_{3,2}(T_3) - c_{2,3}(T_2) \tag{19}$$

$$c_3\dot{T}_3 = c_{2,3}(T_2) + c_{5,3}(T_5) - c_{3,2}(T_3) - c_{3,5}(T_3) - c_{3,4}(T_3 - T_4) \tag{20}$$

$$c_4\dot{T}_4 = c_{3,4}(T_3 - T_4) - c_{4,0}(T_4 - T_0) \tag{21}$$

$$c_5\dot{T}_5 = c_{3,5}(T_3) - c_{5,3}(T_3) - c_{5,6}(T_5 - T_6) \tag{22}$$

$$c_6\dot{T}_6 = c_{5,6}(T_5 - T_6) - c_{6,0}(T_6 - T_0) - c_{6,9}(T_6 - T_9) \tag{23}$$

$$c_7\dot{T}_7 = c_{8,7}(T_8 - T_7) + c_{1,7}(T_1 - T_7) - c_{7,0}(T_7 - T_0) \tag{24}$$

$$c_8\dot{T}_8 = c_{1,8}(T_1 - T_8) + \chi c_{9,8}(T_9) + (1-\chi)c_{13,8}(T_{13}) - c_{8,7}(T_8 - T_7) - c_{8,13}(T_8) \tag{25}$$

$$c_9\dot{T}_9 = c_{6,9}(T_6 - T_9) + \chi c_{10,9}(T_{10}) - \chi c_{9,8}(T_9) \tag{26}$$

$$c_{10}\dot{T}_{10} = \chi c_{13,10}(T_{13}) - \chi c_{10,9}(T_{10}) - c_{10,11}(T_{10} - T_{11}) \tag{27}$$

$$c_{11}\dot{T}_{11} = c_{10,11}(T_{10} - T_{11}) - c_{11,12}(T_{11} - T_{12}) \tag{28}$$

$$c_{12}\dot{T}_{12} = c_{11,12}(T_{11} - T_{12}) + c_{0,12}(T_0) - c_{12,0}(T_{12}) \tag{29}$$

$$c_{13}\dot{T}_{13} = c_{14,13}(T_{14} - T_{13}) + c_{8,13}(T_8) - \chi c_{13,10}(T_{13}) - (1-\chi)c_{13,8}(T_{13}) \tag{30}$$

$$c_{14}\dot{T}_{14} = c_{15,14}(T_{15} - T_{14}) - c_{14,13}(T_{14} - T_{13}) - c_{14,0}(T_{14} - T_0) \tag{31}$$

$$c_{15}\dot{T}_{15} = c_{16,15}(T_{16}) - c_{15,16}(T_{15}) - c_{15,14}(T_{15} - T_{14}) \tag{32}$$

$$c_{16}\dot{T}_{16} = c_{dost2} + c_{15,16}(T_{15}) - c_{16,15}(T_{16}) - c_{16,17}(T_{16} - T_{17}) \tag{33}$$

$$c_{17}\dot{T}_{17} = c_{16,17}(T_{16} - T_{17}) - c_{17,0}(T_{17} - T_0) \tag{34}$$

The values of c_i and c_{ij} factors were estimated on the basis of technical documentation of individual elements of the powertrain. The c_i

values describe the unit thermal capacities of the distinguished elements and equal the product of the mass of the element [kg] and its specific heat [kJ/kg·K]. The c_{ij} values describe the heat transfer conditions on the surface of the mentioned element and are equal to the product of the heat transfer surface of the element [m^2] and the heat transfer coefficient on the surface [kJ/m^2·K] (see Table 1).

The model of heating the engine itself without a retarder has been verified experimentally on the engine dynamometer in the Laboratory of Combustion Engines of the Wroclaw University of Technology and Science (Fig. 5). The results obtained from the measurement of the selected parameters describing the operation of the cooling system were used for determining the threshold conditions in the simulation model. Matlab-Simulink package was used for simulation tests (Fig. 6).

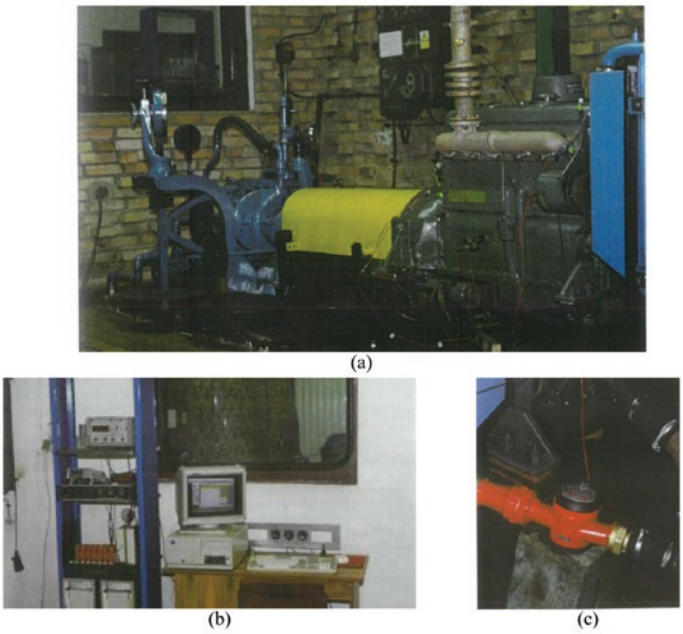


Fig. 5. Pictures of the test stand: (a) view of the 6C107 engine on the engine dynamometer of the Wroclaw University of Technology and Science, (b) computerized test bench on a dynamometer, with a temperature measurement module, (c) view of the engine block coolant flow meter

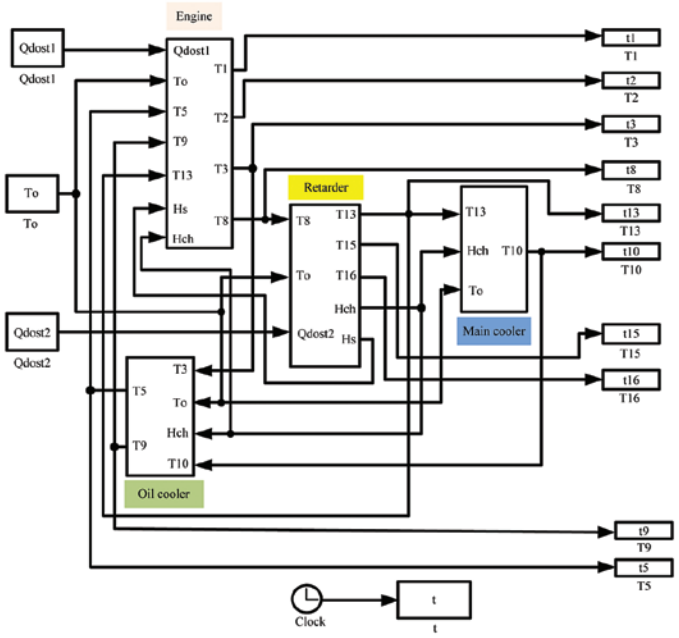


Fig. 6. Matlab-Simulink model

4. Results of simulation tests

Simulation tests were preceded by the development of a test plan. All components of the powertrain, i.e. temperatures from T_1 to T_{16}, distinguished in the temperature model, according to the legend to Fig. 6. A series of tests was carried out taking into account different ambient temperatures, heat exchange surfaces of the radiator, increasing the thermal capacity of the system. The tests used a cooling liquid based on:

a) water-filled engine cooling system,
b) glycol fluid-filled cooling system.

The continuous line indicates temperature fluctuations for the analysed subsystems – T_3, T_8 and T_{16}. The dotted line indicates threshold values in the analysed subsystems.

4.1. Simulation tests of a cooling system at different ambient temperatures

The cooling system tests were carried out at ambient temperatures (T_{ot}) in the range of 253K to 313K every 10K steps. Examples of results are shown in Fig. 7 and Fig. 8. According to them, the system can operate safely in the ambient temperature range up to T_{ot} = 306K (33ºC) (the highest ambient temperatures in Poland reach 313K (40ºC). The improvement of the situation can be achieved by enlarging e.g. the cooler of the engine block coolant.

4.2. Simulation tests of the cooling system for the case of an increase in the heat exchange area of the main cooler

The tests were carried out for three variants of cooler sizes: increased by 10%, by 30% and by 50%, at ambient temperature T_{ot} = 313K (40ºC). The results are presented in Fig. 9.

As can be seen from the presented results, in a glycol-filled system, the use of a cooler with capacity increased by 50% effectively protects all media against exceeding the permissible temperatures, especially the temperature of oil in the retarder. The use of a 50% larger main cooler led to a reduction in oil temperature of about 12K. However, this is not an optimal solution. It is worth investigating the impact that increasing the capacity of the pump in the main cooling system can have on overall capacity.

4.3. Simulation tests of the vehicle cooling system for the case of the cooler increased by 50% and increased coolant flow

The studies were carried out, as in previous cases, for water and glycol at increased flows by 50% and by 100% at ambient temperature T_{ot} = 313K (40ºC). The results are shown in Fig. 10.

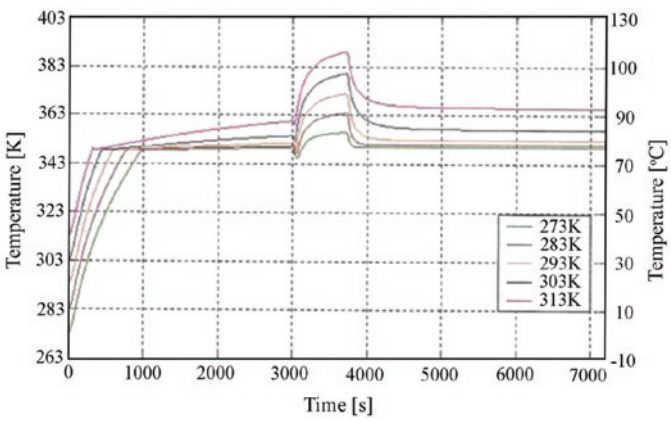


Fig. 7. Trend of cooling liquid temperature changes in engine block at different ambient temperatures (experiment)

Table 1. List of coefficients describing the engine thermal state

No.	A	B	C	D
1	c_i [J/K]	$c_{i,j}$ [W/K]	$c_{i,0}$ [J/K]	other parameters
2	$c_1 = 87990$	$c_{1,2} = 16$	$c_{4,0} = 1.2$	$T_0 = 293.15$ [K]
3	$c_2 = 11750$	$c_{1,7} = 200$	$c_{6,0} = 10$	$q_{dost1} = 29500$ [W]
4	$c_3 = 1796$	$c_{1,8} = 998$	$c_{7,0} = 60$	$q_{dost2} = 91263$ [W]
5	$c_4 = 5560$	$c_{2,3} = 20$	$c_{12,0} = 13000$	$T_p = 348$ [K]
6	$c_5 = 6080$	$c_{3,4} = 097$	$c_{14,0} = 12.5$	$T_k = 358$ [K]
7	$c_6 = 1800$	$c_{3,5} = 190$	$c_{17,0} = 50$	-
8	$c_7 = 448148$	$c_{5,6} = 72$	-	-
9	$c_8 = 20000$	$c_{6,9} = 100$	-	-
10	$c_9 = 20950$	$c_{8,7} = 236$	-	-
11	$c_{10} = 41900$	$c_{9,8} = 5866$	-	-
12	$c_{11} = 19250$	$c_{9,10} = 5866$	-	-
13	$c_{12} = 91$	$c_{10,11} = 955.6$	-	-
14	$c_{13} = 20960$	$c_{11,12} = 1240$	-	-
15	$c_{14} = 4500$	$c_{13,18} = 5866$	-	-
16	$c_{15} = 6840$	$c_{13,10} = 5866$	-	-
17	$c_{16} = 3040$	$c_{14,13} = 583$	-	-
18	$c_{17} = 24750$	$c_{15,14} = 456$	-	-
19	-	$c_{16,15} = 5866$	-	-
20	-	$c_{16,17} = 847$	-	-

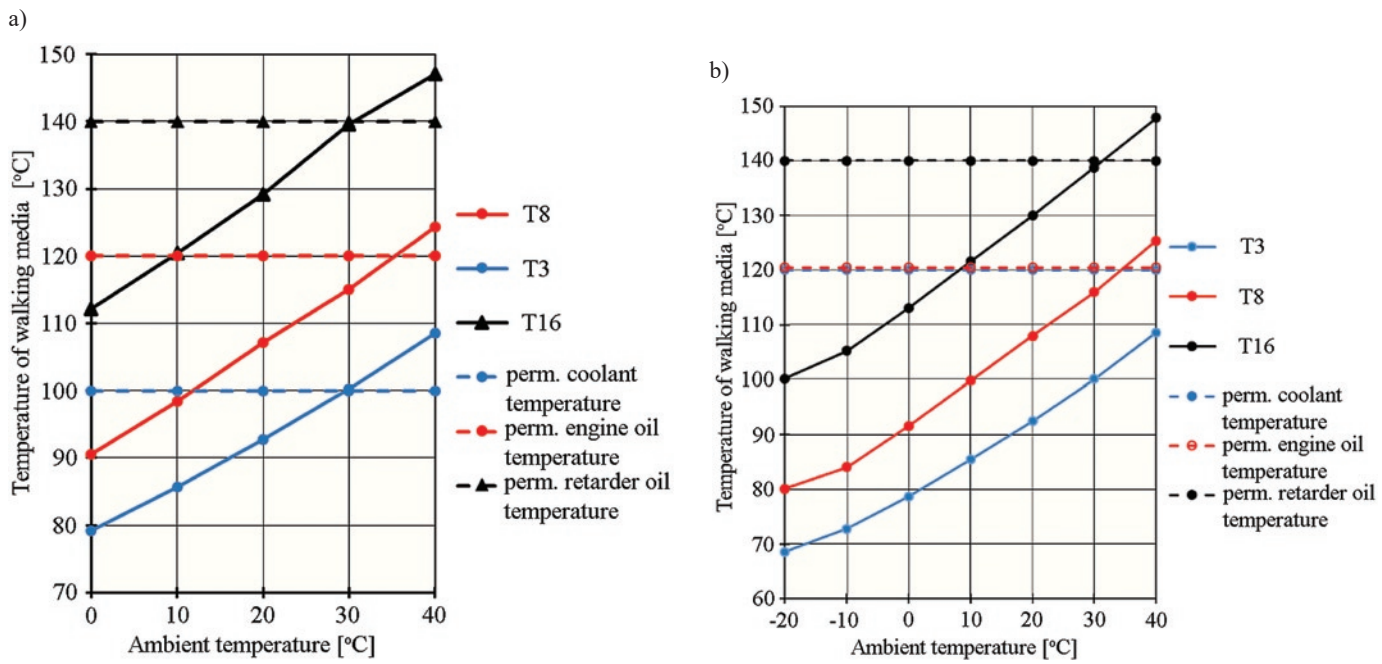


Fig. 8. Temperature of cooling media T_8, T_3, T_{16} versus ambient temperature T_{ot} (experiment): (a) for water, (b) for glycol

Research results show that an effective solution to the problem of temperature reduction in the retarder cooling system can be to increase the pump capacity by 50% or even only by 25% [15, 16].

4.4. Simulation tests of the vehicle's cooling system when an additional coolant tank is used to increase the thermal capacity of the system

The tests were carried out for tanks with a capacity of 20 and 40 dm^3 at an ambient temperature of T_{ot} = 313K (40ºC). The tank was located in the so-called large cooling circuit between the retarder cooler and the main engine cooler. It was assumed that the cooling system is filled with water. The results are shown in Fig. 11.

Equipping the cooling system with an additional coolant tank with a capacity of up to 40 dm^3 causes a slight decrease in system temperature. This solution is ineffective.

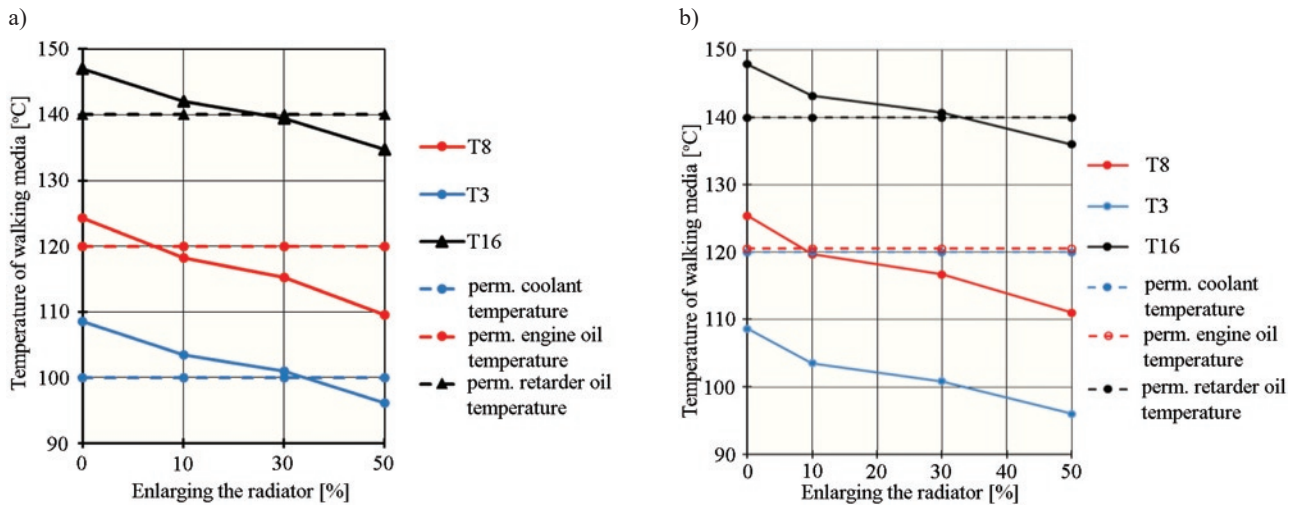


Fig. 9. Temperature of cooling media T_8, T_3, T_{16} versus the size of the radiator T_{ot} (simulation): (a) for water, (b) for glycol

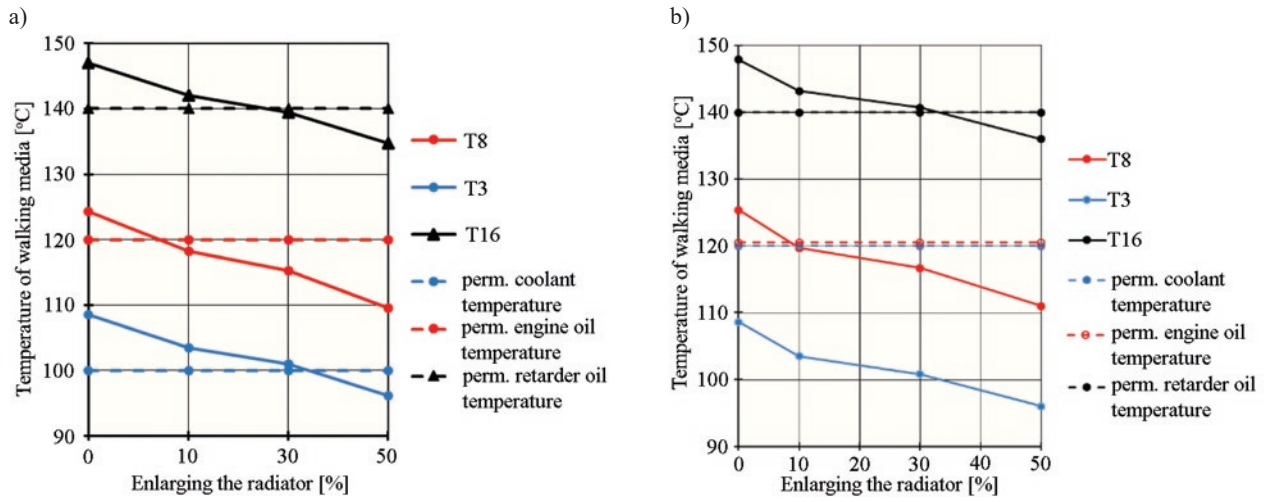


Fig. 10. Temperature of cooling media T_8, T_3, T_{16} versus coolant flow T_{ot} (simulation): (a) for water, (b) for glycol

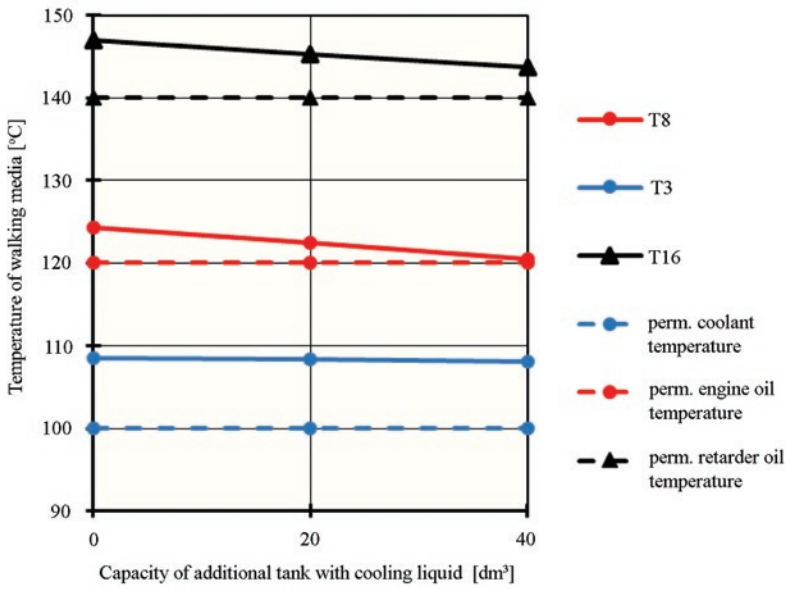


Fig. 11. The dependence of cooling media temperatures on the size of the additional coolant tank (simulation)

5. Conclusion

The aim of the study was to analyze the dynamics of heat exchange in a vehicle equipped with a hydraulic retarder. This device emits large energy flows into the cooling system of the braked vehicle. Therefore, the cooling system should take into account the retarder operation. The requirements in this respect are defined by the relevant EC regulations. On the basis of the analysis of the cooling system structure, a structural and computational model was built. The model of the cooling system is based on a set of 17 balance equations which, according to Newton's principle, were transformed into a set of 17 differential equations describing temperature changes of the distinguished elements versus time. The Matlab-Simulink package was used to solve the equation system. In the study, analyses of the influence of working conditions and constructional conditions on the behaviour of the system in the assumed scenario of its operation were carried out. The results of the work allowed to indicate the direction in which the design of the cooling system of the drive unit should be upgraded so that it could perform its function even in the most difficult working conditions.

References

1. Abe K, Kondoh T, Fukumura K, Kojima M. Three-dimensional simulation of the flow in a torque converter. SAE 1991, https://doi.org/10.4271/910800.
2. Albertz D, Dappen S, Henneberger G. Calculation of the 3D nonlinear eddy current field in moving conductors and its application to braking systems. IEEE Transactions on Magnetics 1996; 32(3): 768-771, https://doi.org/10.1109/20.497353.
3. Baranowski P, Damaziak K, Malachowski J. Brake system studies using numerical methods. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(4): 337-341.
4. Baranowski P, Damaziak K, Malachowski J, Mazurkiewicz L, Kastek M, Polakowski H, Piatkowski T. Experimental and numerical tests of thermomechanical processes occurring on brake pad lining surface. Surface Effects and Contact Mechanics 2011; 10: 15-24, https://doi.org/10.2495/SECM110021.
5. Christoffersen S, Wallingford J, Greenlees B. Heavy truck engine retarders. Testing and theory. SAE 2011, https://doi.org/10.4271/2011-01-0280.
6. Dong Y, Korivi V, Attibele P, Yuan Y. Torque converter CFD engineering part I: Torque ratio and K factor improvement through stator modifications. SAE 2002, https://doi.org/10.4271/2002-01-0883.
7. ECE Regulation No. 13, Uniform provisions concerning the approval of vehicles of categories M, N and O with regard to braking.
8. Gohring E, Glasner EC, Povel R. Engine braking systems and retarders - an overview from European standpoint. SAE 1992, https://doi.org/10.4271/922451.
9. Habib G. The present status of electro-magnetic retarders in commercial vehicles. SAE 1992, https://doi.org/10.4271/922450.
10. Haiss G. Demand criteria on retarders. SAE 1992; https://doi.org/10.4271/922453.
11. Heisler H. Advanced Vehicle Technology. Second Edition. Butterworth-Heinemann: Elsevier Ltd., 2002.
12. Jeyakumar S, Sasikumar M. Computational fluid dynamics simulation of hydraulic torque converter for performance characteristics prediction. International Journal of Scientific Research in Science, Engineering and Technology 2017; 3(6): 402-408.
13. Jia Y H. Dynamic simulation research of hydrodynamic retarder in brake process. Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE) 2011: 1193-1196.
14. Kazmierczak A, Krakowian K, Wrobel. Doppler Laser Vibrometry in combustion engine's diagnosis. Przeglad Elektroniczny 2010; 86(10): 147-149.
15. Kern J, Ambros P. Concepts for a controlled optimized vehicle engine cooling system. SAE 1997, https://doi.org/10.4271/971816.
16. Kesy A, Kadziela A. Construction optimization of hydrodynamic torque converter with application of genetic algorithm. Archives of Civil and Mechanical Engineering 2011; 11(4): 905-920, https://doi.org/10.1016/S1644-9665(12)60086-7.
17. Kwasniowski S, Sroka Z. Modeling dynamics of heat exchange in cooling, heating and air conditioning systems of vehicles and working machines. Wroclaw: Series report SPR No. 075/97, Wroclaw University of Technology, 1998.
18. Lei Y, Song P, Zheng H, Fu Y, Li X, Song B. Application of fuzzy logic in constant speed control of hydraulic retarder. Advances in Mechanical Engineering 2017; 9(2): 1-11, https://doi.org/10.1177/1687814017690956.
19. Li J, Tan G, Ji Y, Zhou Y, Liu Z, Xu Y. Design and simulation analysis for an integrated energy-recuperation retarder. SAE Technical Paper 2016, https://doi.org/10.4271/2016-01-0458.
20. Li R, Yang J, Zhang W. Simulation Study of the Vehicle Hydraulic Retarder, International Journal of Control and Automation 2015; 8(2): 263-280, https://doi.org/10.14257/ijca.2015.8.2.26.
21. Liu C Y, Jiang K J, Zhang Y. Design and use of an eddy current retarder in an automobile. International Journal of Automotive Technology 2011; 12(4): 611-616, https://doi.org/10.1007/s12239-011-0071-3.
22. Marechal Y, Meunier G. Computation of 2D and 3D eddy currents in moving conductors of electromagnetic retarders. IEEE Transactions on Magnetics1990; 26(5): 2382-2384, https://doi.org/10.1109/20.104738.
23. Mu H, Yan Q, Wei W. Study on influence of inlet and outlet flow rates on oil pressures and braking torque in a hydrodynamic retarder. International Journal of Numerical Methods for Heat & Fluid Flow 2017; 27(11): 2544-2564, https://doi.org/10.1108/HFF-10-2016-0428.
24. Pandey S N, Khaliq A, Zaka M Z, Saleem M S, Afzal M. Retarder used as braking system in heavy vehicles - a review. International Journal Mechanical Engineering and Robotic Research 2015; 4(2): 86-90.
25. Peng Z. A Study into the technology of development of hydraulic. Xi'an: Chang'an University, 2008.
26. Pernestål A, Nyberg M, Warnquist, H. Modeling and inference for troubleshooting with interventions applied to a heavy truck auxiliary braking system. Engineering Applications of Artificial Intelligence 2012; 25(4): 705-719, https://doi.org/10.1016/j.engappai.2011.02.018.
27. Song B, Lv J, Liu Y, Kong F. The simulation and analysis on engine and hydraulic retarder continual braking performance of the tracked vehicle on long downhill. Proceedings 9th International Conference on Electronic Measurement & Instruments 2009: 3-928-3-931, https://doi.org/10.1109/ICEMI.2009.5274169.
28. Sarkar S, Rathod P P. Review paper on thermal analysis of ventilated disc brake by varying design parameters. International Journal of Engineering Research & Technology 2013; 2(12): 1077-1081.
29. Tan G, Guo X, Yang T. Simulation based heavy truck driveline components thermal analysis system. The 9th International Conference on, Electronic Measurement & Instruments, ICEMI'2009: 4-675-4-680, https://doi.org/10.1109/ICEMI.2009.5274676.
30. Tan G, Guo X. The modeling and performance analysis of the retarder thermal management system. SAE 2012, https://doi.org/10.4271/2012-01-1929.
31. Wambsganss M W. Thermal management in heavy vehicles: a review identifying issues and research requirements. Argonne National Lab., IL (US), No. ANL/ET/CP-98208, 1999.
32. Wangand G, Shan S. Review of meta modeling techniques in support of engineering design optimization. Journal of Mechanical Design, Transactions of the ASME 2007; 129(4): 370-380, https://doi.org/10.1115/1.2429697.
33. Wrzecioniarz P, Kwasniowski S, Jamroziak K. Criterion for choosing the power unit cooling system of evacuation tractor. Czasopismo Techniczne Mechanika 1998; 95(5-M): 83-93.
34. Wrzecioniarz P, Kwasniowski S, Jamroziak K. The concept of the cooling system for the road tractor unit. Pojazdy samochodowe: problemy

rozwoju jakości, VI Międzynarodowa Konferencja Naukowo-Techniczna "Autoprogres'98" 1998; T.2: 127-134.
35. Xin Q. Diesel engine system design. New Delhi: Woodhead Publishing, 2011, https://doi.org/10.1533/9780857090836.
36. Yang J, Yi F, Wang J. Model-based adaptive control of eddy current retarder. Proceeding of the 30th Chinese Control And Decision Conference (2018 CCDC) 2018; 1889-1891, https://doi.org/10.1109/CCDC.2018.8407434.
37. Zheng H, Lei Y, Song P. Design of a filling ratio observer for ahydraulic retarder: An analysis ofvehicle thermal management and dynamic braking system. Advances in Mechanical Engineering 2016; 8(10): 1-8, https://doi.org/10.1177/1687814016674098.
38. Zheng H, Lei Y, Song P. Hydraulic retarders for heavy vehicles: Analysis of fluid mechanics and computational fluid dynamics on braking torque and temperature rise. International Journal of Automotive Technology 2017; 18(3): 387-396, https://doi.org/10.1007/s12239-017-0039-z.
39. Zhou L, Tan G, Guo X, Chen M, Ji K, Li Z. Yang Z. Study of energy recovery system based on organic rankine cycle for hydraulic retarder. SAE 2016, https://doi.org/10.4271/2016-01-0239.

Krzysztof JAMROZIAK
Stanislaw KWASNIOWSKI
Mariusz KOSOBUDZKI
Pawel ZAJAC
Faculty of Mechanical Engineering,
Wroclaw University of Science and Technology
ul. Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland

E-mails: krzysztof.jamroziak@pwr.edu.pl, stanislaw.kwasniowski@pwr.edu.pl, mariusz.kosobudzki@pwr.edu.pl, pawel.zajac@pwr.edu.pl

Article citation info:
CASTILLA-GUTIÉRREZ J, FORTES JC, PULIDO-CALVO I. Analysis, evaluation and monitoring of the characteristic frequencies of pneumatic drive unit and its bearing through their corresponding frequency spectra and spectral density. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 585–591, http://dx.doi.org/10.17531/ein.2019.4.7.

Javier CASTILLA-GUTIÉRREZ
Juan Carlos FORTES
Inmaculada PULIDO-CALVO

ANALYSIS, EVALUATION AND MONITORING OF THE CHARACTERISTIC FREQUENCIES OF PNEUMATIC DRIVE UNIT AND ITS BEARING THROUGH THEIR CORRESPONDING FREQUENCY SPECTRA AND SPECTRAL DENSITY

ANALIZA, OCENA I MONITOROWANIE CZĘSTOTLIWOŚCI CHARAKTERYSTYCZNYCH PNEUMATYCZNEGO ZESPOŁU NAPĘDOWEGO I WCHODZĄCEGO W JEGO SKŁAD ŁOŻYSKA NA PODSTAWIE WIDM CZĘSTOTLIWOŚCI ORAZ GĘSTOŚCI WIDMOWEJ

This article shows the results of the study of the characteristic frequencies of pneumatic drive equipment and its suspension bearing. The analysis approaches one of the most important requirements of the industrial sector, which seeks to be recognised by the efficiency and performance of its equipment when compared to its coming economic competitors. For data collection and we have followed the ISO 10816 standards, thus using the values of speed in RMS, aiming to reduce the masking of these signals that occurs depending on whether they are high or low frequencies. The study will respond to one of the most important requirements found in the predictive and preventive control of industrial sites. The problem of the predictive systems of maintenance of equipment with bearings lies in the number of monitoring and analysis points that generate a high cost in time and human resources. The aim will be to determine which of all the study frequencies is the most significant and in which position and measurement axis has the biggest impact. To do this, we will analyse the rotation frequency of the blowing machine, the resulting frequency of all the frequencies, the frequency of the impulsion blades and finally the frequency of the bearing. The study would be able to predict when our equipment is going to suffer a failure, reducing the control points and the cost.

***Keywords**: vibration, bearing, diagnostics, vibroacoustics.*

W artykule przedstawiono wyniki badań częstotliwości charakterystycznych napędu pneumatycznego i wchodzącego w jego skład łożyska zawieszenia. Analiza przybliża jedno z najważniejszych wymagań sektora przemysłowego, w którym dąży się do tego by wyróżniać się na tle konkurencji sprawnością i wydajnością urządzeń. Przy zbieraniu danych postępowaliśmy zgodnie ze normą ISO 10816, wykorzystując średnie prędkości kwadratowe, co pozwoliło zmniejszyć maskowanie sygnałów, które występuje w zależności od tego, czy mamy do czynienia z wysokimi czy niskimi częstotliwościami. Badanie stanowi odpowiedź na jeden z najważniejszych wymogów w zakresie kontroli predykcyjnej i prewencyjnej obiektów przemysłowych. Problemem systemów konserwacji predykcyjnej sprzętu, w którego skład wchodzą łożyska jest duża ilość punktów kontrolnych, które generują wysokie koszty jeśli chodzi o czas i zasoby ludzkie. Celem pracy było określenie, które ze wszystkich badanych częstotliwości są najistotniejsze oraz dla których częstotliwości pozycja i oś pomiaru mają największe znaczenie. W tym celu przeanalizowano częstotliwość obrotową analizowanej dmuchawy, częstotliwość wynikową wszystkich częstotliwości, częstotliwość łopatek oraz częstotliwość łożyska.

***Słowa kluczowe**: drgania, łożysko, diagnostyka, wibroakustyka.*

1. Introduction

In an increasingly globalised world, cost reduction is a strategic and competitive advantage. For this reason, a demand has arisen in the industrial sector for automated and computerised procedures to efficiently control and manage systems based on preventive and predictive maintenance, against corrective actions that generate higher costs in terms of time, resources, material and production losses.

Rotating elements, including bearings, are one of the main causes of shutdown or failure in production systems. Almost 40% of all shutdowns are caused by these rotating elements [11], hence the importance of incorporating control and monitoring systems to predict their service life and evaluate their condition. A bearing's behaviour depends on many variables such as lubrication operations [20], the strengths of excessive or external dynamic loads, design errors, pollution, handling defects and transport shocks.

The conventional procedure of failure diagnosis and analysis of bearings is performed by means of vibration signals, which have a fundamental implication in this type of maintenance. The information provided by vibrations in the frequency domain gives us the keys to determine possible erratic conditions in rotating elements. The most significant ones found in rotating systems are unbalance, misalignment, eccentricity, bent shafts, cavitation in turbo machinery and bearing failures.

The main purpose of vibration monitoring and control is to determine one of the four stages of failure in which the bearing is. It should be

noted that bearings do not show any signs of breakage or anomalies until they reach one fifth of their service life [17]. In the first stage, erratic behaviours are generated in elements whose working frequency is mainly higher than 5kHz. At this stage it is not necessary to replace the bearing and although cracks occur, they are not noticeable to the human eye.

The second stage starts when average values are reached. Then, cracks appear, they are now visible to the human eye can also be perceived acoustically due to the generation of small disturbances. The third and fourth stages are the most important ones. Cracks advance, increasing the temperature of elements, causing their rapid advance towards breakage. At this stage, acoustic perception is very severe and easily perceived to the human ear. The fourth stage involves a corrective action, generating important actions caused by the breakage of components and according to the moment of shutdown of the whole equipment until the total breakage [2].

The first research works based on the use of vibration as a preventive maintenance method for rotating equipment were focused on improving the quality of vibration signal detection systems and their processing to improve the analysis of frequency spectra according to the type of machines (turbines, pumps, fans, gearboxes, compressors) and the type of defects (loads, tensions, misalignment, unbalance, lubrication) [6].

The main goal of new working trends is the prediction of failures as a fundamental basis for preventive maintenance [13]. The acoustic analyses for the characterisation and assessment of the vibrations generated are particularly remarkable. The advantage of these techniques is that they are non-intrusive, since they do not disturb the normal operation of mechanical systems, but daily monitoring of equipment, with the consequent need for human and material resources is needed. This factor conditions the efficiency and productivity of many industries subject to tight profit margins [4].

The most important lines of research are the behaviour of bearings, their friction, the internal stresses and how the friction of their own components affects and the dissipation of energy through the roughness. This study also includes the generation of new designs with gas application to reduce the tangential effects of centrifugal forces for rotating equipment. It is important to note that, all the tests have been carried out in the laboratory without being able to study the evolution over the years in real operating conditions [23].

For this reason, this work evaluates the characteristic operating frequencies of a set of bearings in a blower of a real industrial plant, through the spectral analysis of the vibrations generated during 15 years of operation. Most of the studies carried out on neural networks on the behaviour and durability or useful life of rotation systems, including bearings, have been based on experimental work in the laboratory, ending in their breakage [19, 8].

Other experimental work in laboratory using Support Vector Machine (SVM) focuses on the diagnosis of ball bearing failures by processing the signal through the wavelet transform [12].

2. Material and methods

2.1. Equipment under study

The drive equipment under study is a blower that belongs to the production system of a Spanish industrial plant whose main objective is safety, and of course operational and energy efficiency and environmental protection [3]. Safety has led this company to develop a highly specialized equipment control and monitoring program, seeking for durability and fault prediction. To carry out this process, they have instruments and specialized personnel that sample the equipment daily in different positions, studying its variations and trying to predict its failure. That is why they have facilitated the 15-year sampling of a blower, to try to find a relationship or linearity between the different parameters that affect the operation of the bearings.

The equipment under study is a blower designed for pumping fluids. The equipment consists of two different parts, the impulsion system or motor frame size 400 of 3190 kg, with dimensions 1900 x 910 mm, fixed to a steel bench of 2100 x 3900 x 300 mm. At the same time, it serves as support and connection to the system, formed by the drive unit and a fan of 2075 mm in diameter, all of which is connected by a 95 mm and 3900 mm shaft. A coupler drives the fan blades, so the shaft is divided into two sections, the 1900 mm motor part and the 2000 mm fan part. This section is supported by three bearings: SKF6322, FAGNU322 and SKFNU322 [2].

This equipment, like most modern machines operating at high speeds and loads, has a bearing-supported shaft as a mechanical transmission element. Bearings are the first elements to cause system failure. That is why it is so important to have a proper design, an appropriate selection of bearings and a comprehensive testing plan. The vibration signals from bearings are very complex, because they hide and couple together, and are difficult to analyse under normal and extreme operating conditions.

Many studies of bearing vibration analysis have been conducted, trying to respond to this need of the industrial sector by means of mathematical analysis of vibrations, generating models to predict defects, but none have been conclusive [9]. The problem is that the system is in a dynamic state of motion. This energy is partly absorbed by the material itself, so the rigidity of the rolling element is a fundamental variable for the study of faults.

In these models, the vibration signal of a defect is established through the series of pulses. Hitting the moving parts on the coincident surface, generating resonances in the bearing and housing structures, generates these. This knocking is caused by the system being in continuous motion, when it is in stable operation and rotating at constant speeds. These pulses generate periodic disturbances and can thus determine their frequency, depending on their position [18]. The geometry (rolling elements, cage and races) and running speed of bearings precisely determine their characteristic frequencies [15, 9, 1].

Table 1. Nominal characteristics of the bearing supporting the blower shaft under study

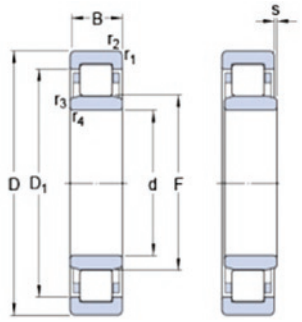
Bearing FAGNU322		
Basic dynamic load (C)	415 kN	
Basic static load (C0)	475 kN	
Fatigue limit load (Pu)	61 kN	
Reference speed	3000 r/min	
Speed limit	3000 r/min	
Calculation factor (kr)	0.15	
Geometric characteristics	d = 110 mm; D = 240 mm; B = 50 mm; $d_1 \approx$ 200 mm; F = 143 mm; $r_{1,2}$ min. 3 mm	

Table 2. Characteristic frequencies of the bearing under study (Hz)

Frequencies	FAGNU322
FTF	508.81
BSF	3938.1
2BSF	7876.1
BPOR	7911.9
BPIR	11452
Speed	1490
Blades Fpa	13410

Each bearing component has a unique breaking frequency and four fundamental frequencies are defined [4], according to the place where the defect occurs. So, if it is generated in the balls we will have the BSF; if it is by the frequency of cage rotation the FTF; if it is by the passage in the inner race, the BPIR and if it is by the outer race BPOR [21].

The equipment under study consists of three different types of bearings. The nominal characteristics of FAGNU322 bearing are shown in Table 1. The fundamental frequencies of the blower shaft bearing are as shown in Table 2. They are compliant with the instructions of ISO 10816.3 [22].

In this study, we have considered not only the characteristic breaking frequencies of the different parts of a bearing, but also the rotation frequency of the system, which in our case has a continuous motion regime without accelerations, generating a stable working motion at a constant angular speed. Another factor to be noted is the characteristic frequency of the driven elements, in this case, the air discharge blades.

2.2. Vibration analysis

Predictive analysis of rotating equipment is possible by transforming the signals caused by its own vibrations. These signals are processed in the frequency domain using the FFT (Fast Fourier Transform) algorithm to relate the shape of the waves generated by vibration to the frequency generating its own vibratory spectrum. Decomposing the vibration, or rather, the wave generated into simple waves according to its characteristic frequency and representing each of them independently, we obtain the representation in the frequency domain of this disturbance; this representation is called spectrum or spectral density.

Signal analysis in the frequency domain creates major impediments to perform a reliable signal analysis. Therefore, the concept of signal energy through the Power Spectral Density (PSD) arises, defined as the amount of energy contained in each characteristic frequency and has the following expression:

$$PSD\left(x(i)\right)=\frac{1}{T}\Sigma\left(x(i)\right)^2\Delta T \tag{7}$$

where the x (i) sequence for ΔT is the average power of the interval or the sum of the powers of each of the components, independently. The simplest way to estimate this power density is to pass the data from a defined sample window. The defined window is called the periodogram and the advantage of this method is that it can detect hidden frequencies [7].

There are other methods that improve the accuracy of the analysis, one of which is the Hilber HT (Hilbert Transform) transformation that analyses the signal envelope. This allows obtaining an improved PSD and facilitates the sampling in processes that take place at low frequency, modulating precisely the primary signal

The method followed in this work was to transform two functions s (t) and 1/(πt) into a third one, as expressed below:

$$\bar{x}(t)=\frac{1}{\pi}\int_{-\infty}^{\infty}x(u)\frac{1}{t-u}du \tag{8}$$

This method can amplify events occurring at low frequencies, through the components that form the signal envelope, but although the system provides improvements, it also has deficiencies since it is very sensitive to noise [14]. The problem was solved with the incorporation of a data treatment process, providing information on the main study variables in signal processing, such as time and frequency [16]. This process is known as Wavelet Transformation (WT) [10].

3. Data acquisition, processing and analysis

The Entek IRD analyser and the OdysseyEmonitor software have been used to obtain vibration data from the blower under study from May 2000 to April 2015. The hardware consists of a 16-channel interface with a nominal voltage of 24 V at 3.6 mA DC, a multiplexer with 4 high-pass filters, an integrator to generate speed-dependent acceleration signals and, similarly, speed-dependent displacement signals. Another important element is the antialiasing filter that eliminates high frequencies, avoiding the appearance of aliasing in the received signal and, finally, the digital analogue converter that captures up to 51.2 kHz with a resolution of 16 bits. The OdysseyEmonitor Software allows you to generate trend graphs of vibration energy levels and their characteristic frequencies.

According to the risk levels of the ISO 10816-3 standard and the critical operating levels, the equipment under study has a flexible shaft, power higher than 300 kW and a speed of 1,500 rpm, so it would have an operating level between 0.18 to 11 RMS (mm/s), with the maximum critical value being 7.1 RMS, where corrective maintenance must be applied.

The equipment consists of two measuring points identified as 3 and 4; at each point, sampling is performed at the three coordinates of the X, Y, and Z space, identified as shown in (Table 3).

An accelerometer with a sensitivity level of 100 mV/g and a translator with contact displacement were used, allowing a frequency range of 10000 Hz. The fixing system will be a magnet and a fast mounting base, which allows reaching faults in a higher frequency range, improving its range and obtaining measurement sensitivity between 0-300 Hz. The monitoring is done with a Hanning window, 3,200 lines, and a speed range of 60,000 and 300,000 Hz for acceleration, obtaining a bandwidth of 28 CPM for speed and 140 CPM for acceleration and a resolution of 18 CPM for speed and 93 CPM for acceleration.

For faults diagnosis, the ISO 10816 standard determines that the evaluation is performed using the spectral level based on speed. This presents greater uniformity at both low and high frequencies.

Table 3. Sampling positions of the FRL104 blower under study

Identification	Number	Site
Sopl. LA H	003	Horizontal
Sopl. LA V	003	Vertical
Sopl. LA A	003	Axial
Sopl. LOA H	004	Horizontal
Sopl. LOA V	004	Vertical
Sopl. LOA A	004	Axial

In the case of acceleration, better resolution levels are obtained at high frequencies and in the case of displacement it is more representative at low frequencies.

One of the problems found when using acceleration or displacement is the skiing effect, caused by the integration of noise, which can hide defects due to the attenuation of high frequencies and the enhancement of components generated at low frequencies. In this way, digital integration is performed on the vibratory spectrum of the vibration and weights the low frequency signals, acting less on the high frequency signals, this is because the speed V(f) is inversely proportional to the frequency f:

$$V(f)=\frac{c_1\ A(f)}{f} \tag{9}$$

while the displacement D(f) is also affected by the frequency, but in an exponential way:

$$D(f)=\frac{c_2\ A(f)}{f^2} \tag{10}$$

where A(f) is the acceleration to frequency f and C1 and C2 are constant according to the units of measurement.

This study is compliant with the provisions of the standard and the maximum MSY levels are used, obtaining a total of 617 measurements in each of the six positions, having, therefore, a total of 3,702 measurements. An error of 5% has been considered for the sampling of each characteristic frequency and an error of 2.5% for each position. From the three measuring points, only the horizontal position had continuous monitoring, the rest was done in a reduced the rest was done on a reduced basis over periods of between seven and 15 days. Hence, the measurement value obtained at the same time was about 600 measurements in those 15 years for each of the positions.

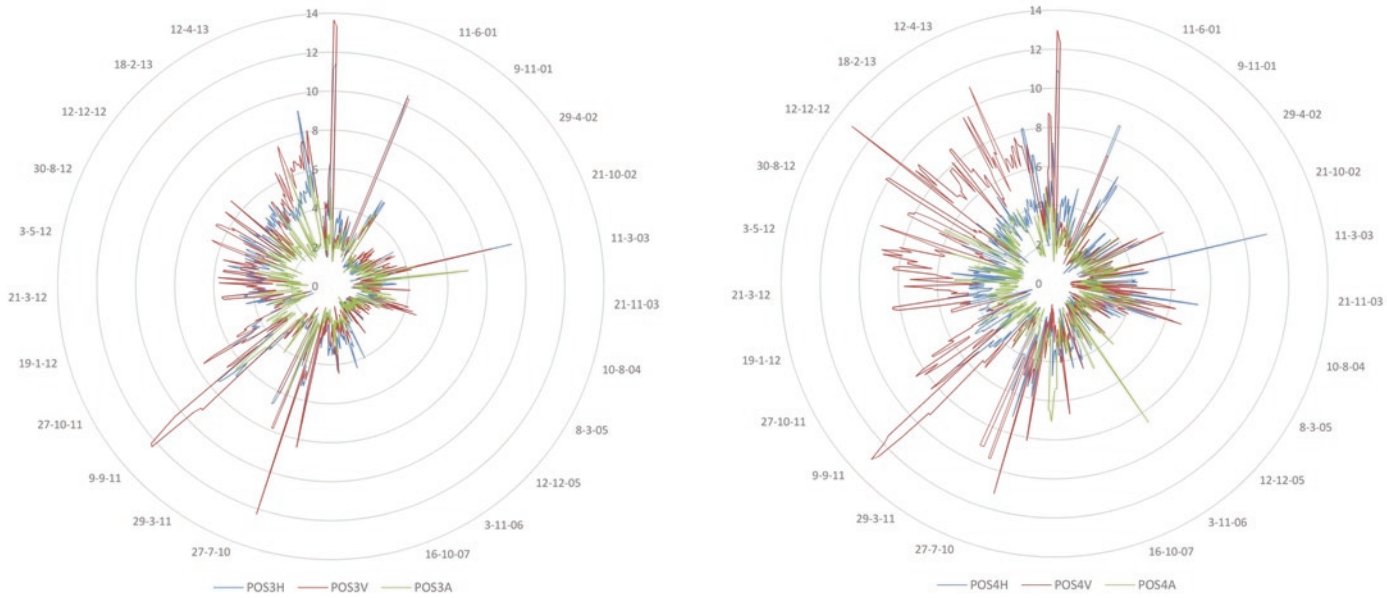


Fig. 1. Graph of Maximum Amplitude results in RMS for positions 3 and 4

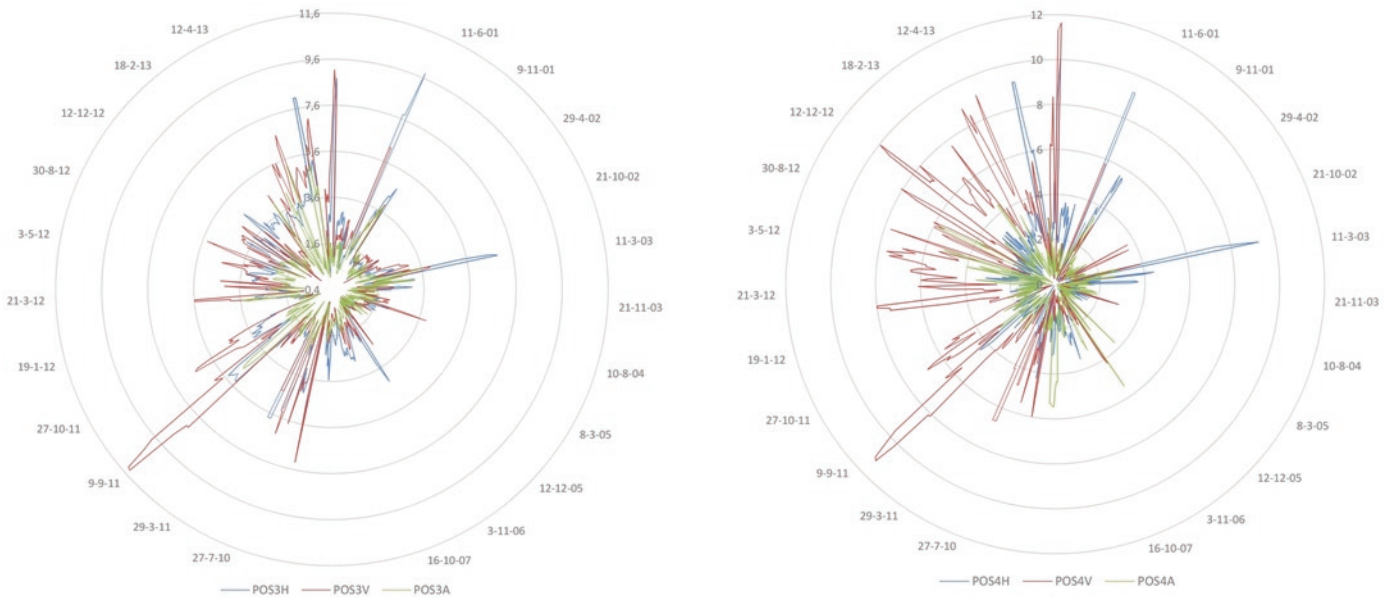


Fig. 2. Graph of SPEED results in RMS for positions 3 and 4

4. Results and discussion

The first study visualizes the maximum amplitude values generated by the digital integration of each spectrum. This analysis is performed in all positions and axes.

The first position of study will be position 3, where the highest average values in the vertical axis (3.46 RMS) are higher than in the horizontal one (3.12 RMS). Finally, the axial axis has an average value in RMS of 2.27. These values indicate that the most incident energy points are in the vertical position.

Another important value is that of peak values. In this case, the maximum value is found in the vertical position with 13.34 RMS, followed by the horizontal one with 11.44 RMS and finally the axial axis with 7.08 RMS. As in the previous case, in position 4 we observe that the maximum value is in the vertical position, followed by the horizontal position, which reaches 84.39% of the value in the vertical axis, and finally the axial with 60.28% of the maximum (4.23 RMS, 3.57 RMS, 2.55 RMS).

In the case of peak values, we found a repetition of the sequence in the previous position and the maximum value is in the vertical axis with 13.12 RMS, followed by the horizontal 11.15 RMS and the last the axial with 8.58 RMS. In position 4 values are much more intense than in position 3. And regarding the axes, they are higher in the vertical axis compared to the rest of positions (horizontal and axial), as shown in (Fig. 1).

After the previous study, we analyse the values of the characteristic frequency of the machine (SPEED), which is at a frequency of 1490 Hz. The first position of study is position 3 and its three axes. For this position the values obtained were: the horizontal axis reaches an average value of 2.17 RMS; in the vertical one the value is 2.48 and in axial, 1.26 RMS.

The results according to the peak values are: 9.82 RMS in the horizontal axis; 11.38 RMS in the vertical axis and 5.32 RMS in the axial axis. It is verified that the vertical axis is again the most sensitive to the rotation frequency of the equipment. After analysing the SPEED frequency in position 3, it is studied in position 4. The study shows the following values: 1.61 in the horizontal axis; 2.61 in the vertical one and 1.34 RMS in the axial. Therefore, the previous results are reproduced again, obtaining the highest energy level in the vertical axis.

According to the peak values, the results are: 10.04 RMS in the horizontal axis; 11.64 RMS in the vertical axis and 5.76 RMS in the axial axis. Once again the vertical axis appears as the most sensitive to the rotation frequency of the equipment, as shown in (Fig. 2).

Next, we study the frequency of the blades. In position 3 the highest RMS value is reached in the axial axis (3.39 RMS), and in the other two axes, horizontal and vertical, the values are 0.18 RMS and 0.13 RMS, respectively. In the case of peak values, the highest value is also found in the axial position with 3.35 RMS; 1.21 RMS in the horizontal axis and 0.97 RMS in the vertical one. Unlike all previous studies where the most relevant value was found in the vertical axis, in this case is in the axial axis where this frequency affects the most. This is due to the effect of impulsion generated by the profile of the blade. The graph also shows that the sequence obtained in position 3 is repeated in position 4, obtaining the highest value in the axial axis with 0.42 RMS, as shown in (Fig. 3).

In the axial axis we obtain an average value of 0.39 RMS for position 4 and 0.42 RMS for position 3. This is the first frequency whose highest value is reached in position 3. Peak results credit the previous result, with 3.35 RMS at position 3 and 1.94 RMS at position 4.

After analysing frequencies and the SPEED of the blade, we carried out the analysis of the frequencies of the FAGNU322 bearing. The most significant variables of bearings are the frequencies generated by the cage of balls defined as FTF, those caused by the balls (BSF), followed by the ones of its second harmonic (2BSF), those generated by the balls in the outer track (BPOR), and the one generated in the inner track (BPIR). After analysing all the significant variables of the FAGNU322 bearing and their influence in each measured axis and position, the results in terms of RMS are shown in the table below, as shown in (Table 4). The positions where the maximum values are generated have been: vertical position 4 for FTF, and 2BSF. The other two, BPOR and BPIR, are in axial axis 3 and in horizontal axis 4, respectively.

Frequencies generate their greatest influence on the vertical axis, followed by the axial one, and being the horizontal axis the least important.

After studying the variables of the FAGNU322 bearing, we compare them with the results of the other variables under study, the rotation frequency of the machine called SPEED and the frequency of the blades. The table of results in terms of RMS is shown in table 5. We can conclude from the results shown in the table that the most important frequency is that caused by the machine itself, much higher than those of the blades and the cage of the bearing or FTF.

When checking the peak values, again we take SPEED with 11.644 RMS as reference, but in this case the second one is not the one generated by the blades but by the frequency FTF FAGNU322, and in the third place that of the blades, as shown in (Fig. 4).

5. Conclusions

The study determines that the most important maximum amplitude values appear in position 4, and that the vertical axis is the most sensitive to the mechanical actions of the equipment. The above result is also valid for specific frequencies, such as the one generated by the rotation frequency of the equipment (SPEED). The analysis of the frequency of the blades shows us that the most determining position

Table 4. Summary of FAGNU322 bearing frequencies in terms of RMS.

VARIABLE Frecuency Position	**FTF** Cage Defects Pos4V	**BSF** Ball Defects Pos4V	**BPOR** Outer Race Defects Pos3A	**BPIR** Inner Race Defects Pos4H	**2 BSF** Second Harmonic, Ball Defects Pos4V
TOTAL SUM	217,471	44,126	38,223	22,819	64,616
AVERAGE	0,355	0,072	0,062	0,037	0,105
PEAK VALUE	10,056	1,392	0,545	0,388	0,659

Table 5. Summary of SPEED, ALABES and FTF frequencies in terms of RMS.

VARIABLE	**SPEED Pos4V**	**FTF Pos4V**	**ALABES Pos3A**
TOTAL SUM	1606,085	217,471	245,778
AVERAGE	2,612	0,355	0,399
PEAK VALUE	11,644	10,056	3,354

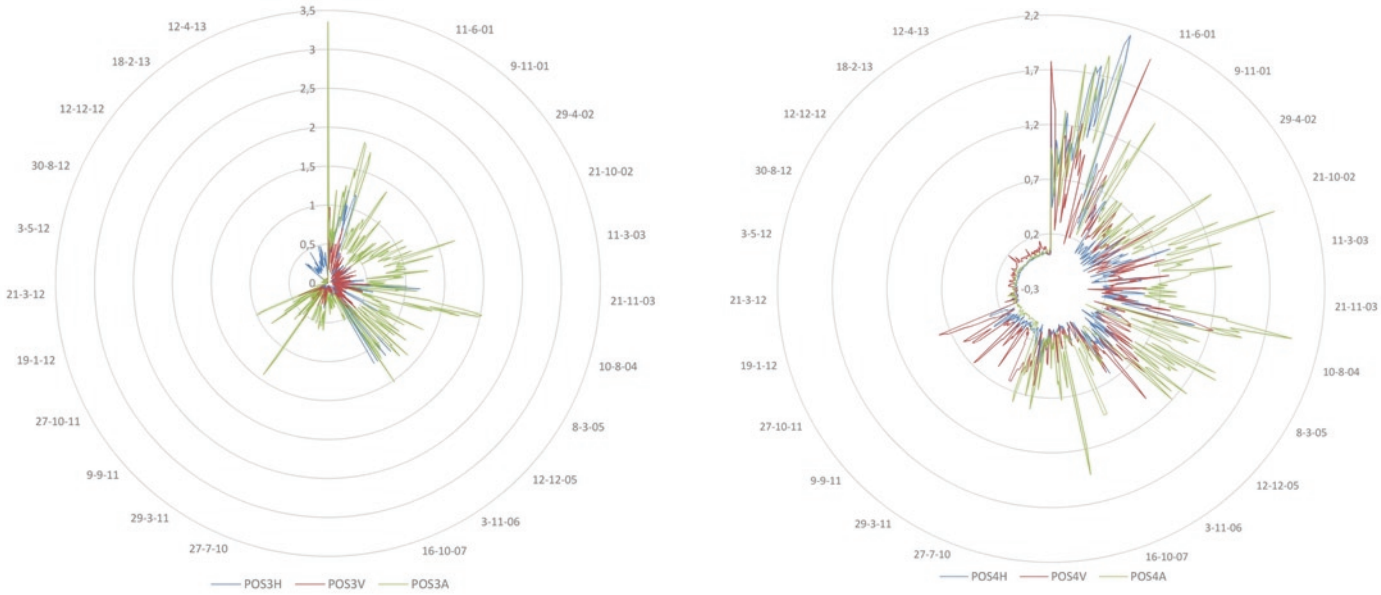


Fig. 3. Graph of the frequency results of blades in RMS for positions 3 and 4

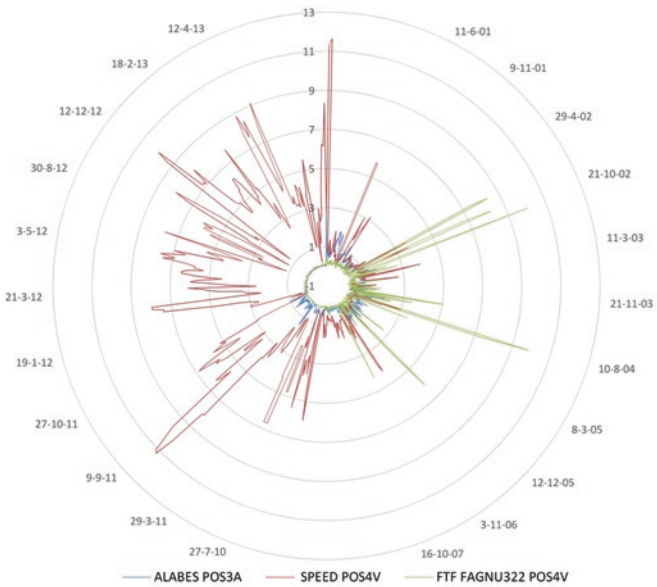


Fig. 4. Comparison of SPEED, ALABES and FTF frequencies FAGNU6322

is position 3 on the axial axis. This is due to the thrust action generated by the blades and the fluid on the axis itself. On the FAGNU6322 bearing, the study determines that the FTF frequency is the most important and reaches its highest value in the vertical axis position 4. The second most important frequency is the one generated by the 2BSF, with average and peak values much lower than the frequency of the FTF. This indicates that the most determining frequency in the operation of the equipment is SPEED in vertical position 4, as well as axial position 3 by the action of the blades in the background. The final conclusion would be that continuously evaluating position 4 on the vertical axis we would be able to predict when our equipment is going to suffer a failure, reducing the control points by 5.

References

1. Artzer A, Moats M, Bender J. Removal of Antimony and Bismuth from Copper Electrorefining Electrolyte: Part I—A Review. JOM 2018, https://doi.org/10.1007/s11837-018-3075-x.
2. Castilla J, Fortes JC, Davila JM, Melgar S, Sarmiento A. Predictive Maintenance of mining machinery based on vibrational analysis. 18 th. International Multidisciplinary Scientific Geoconference & Expo. Sgem 2018, http://doi.10.5593/sgem2018/1.3.
3. Chudzik A, Warda B. Effect of radial internal clearance on the fatigue life of the radial cylindrical roller bearing. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (2): 211–219, http://dx.doi.org/10.17531/ein.2019.2.4.
4. Cong F, Chen G, Dong G, et al. Vibration model of rolling element bearings in a rotor-bearing system for fault diagnosis. J. Sound Vib. 2013; 332 (8): 2081–2097, https://doi.org/10.1016/j.jsv.2012.11.029.
5. Kausschinger B, Schroeder S. Uncertainties in Heat Loss Models of Rolling Bearings of Machine Tools, Procedia CIRP 46 2016; 107 – 110, https://doi.org/10.1016/j.procir.2016.03.168.
6. Leturiondo U, Salgado O, Galar D. Multi-body modelling of rolling element bearings and performance evaluation with localised damage. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18 (4): 638–648, http://dx.doi.org/10.17531/ein.2016.4.20.
7. Li H, Fu L, Zheng H. Bearing fault diagnosis based on amplitude and phase map of Hermitian wavelet transform. Journal of Mechanical Science and Technology 2011; 25(11):2731– 2740, https://doi.org/10.1007/s12206-011-0717-0.
8. McFadden P D, Smith J D. Model for the vibration produced by a single point defect in a rolling element bearing. Journal of Sound and Vibration 1984; 96(1), 69–82, https://doi.org/10.1016/0022-460x(84)90595-9.

9. Medrano-Hurtado Z Y, Medrano-Hurtado C, Pérez-Tello J, Gómez Sarduy M, Vera-Pérez N. Methodology of Fault Diagnosis on Bearings in a Synchronous Machine by Processing Vibro-Acoustic Signals Using Power Spectral Density Ingeniería. Investigación y Tecnología 2016; Volume 17, Issue 1, January–March, Pages 73-85, https://doi.org/10.1016/j.riit.2016.01.007.
10. Mercorelli P, Mercorelli A. Denoising procedure using wavelet packets for instantaneous detection ofpantograph oscillations. Mechanical Systems and Signal Processing 2013; 35(1-2):137–149, https://doi.org/10.1016/j.ymssp.2012.09.001.
11. Nagi G, Alaa E, Jing P. Residual Life prediction sin the absence of prior degradation know ledge. IEEE Trans. Reliab. 2009; (58): 106–116, https://doi.org/10.1109/TR.2008.2011659.
12. Nandi S, Toliyat H A, Li X. Condition Monitoring and Fault Diagnosis of Electrical Motors—A Review. IEEE Transactions on Energy Conversion 2015; 20(4), 719–729, https://doi.org/10.1109/TEC.2005.847955.
13. Patil M S, Mathew J, Rajendrakumar P K, Desai S. A theoretical model to predict the effect of localized defect on vibrations associated with ball bearing. International Journal of Mechanical Sciences 2010; 52(9), 1193–1201, https://doi.org/10.1016/j.ijmecsci.2010.05.005.
14. Pawlik P. Single-number statistical parameters in the assessmente of the techinical condition of machines operating under variable load. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (1): 164-169, http://ds.doi.org/10.17531/ein.2019.1.19.
15. Polimac V, Polimac J. Assessment of present maintenance practices and future trends. IEEE/PES Transmission and Distribution Conference and Exposition. Developing New Perspectives (Cat. No.01CH37294) 2001; https://doi.org/10.1109/tdc.2001.971357.
16. Schnabel S, Marklund P, Larsson R, Golling S. The Detection of Plastic Deformation in Rolling Element Bearings by Acoustic Emission. Tribiology International 2017; https://doi.org/10.1016/j.triboint.2017.02.021
17. Toledo E, Pinhas I, Aravot D, Akselrod S. Bispectrum and bicoherence for the investigation of very high frecuency peaks in heart rate variability. Proceedings of the IEEE, Computers in Cardiology 2001; Núm. 28., pp. 667-670, https://doi.org/10.1109/CIC.2001.977744.
18. Tse P, Peng Y, Yam R. Wavelet analysis and envelope detectionfor rolling element bearing fault diagnosis. Journal of vibration and acoustic 2001; p.303-310, https://doi.org/10.1115/1.1379745.
19. Yujie G, Jinguy L, Jie L, Zhanhui L, Wentao L. A method for improving envelopespectrum symptom of fault rolling bearing based on the auto-correlation acceleration signal. Applied Mechanics and Materials 2013; 275:856–864, https://doi.org/10.4028/www.scientific.net/AMM.275-277.856.
20. Zheng D, Chen W. Thermal performances on angular contact ball bearing of high- speed spindle considering structural constratints under oil-air lubrication. Tribology International, 2017; (109) 593–601, https://doi.org/10.1016/j.triboint.2017.01.035.
21. Zhou W, Habetler T G, Harley R G. Bearing Condition Monitoring Methods for Electric Machines: A General Review. IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives 2007; https://doi.org/10.1109/demped.2007.4393062.
22. Zuber N, Bajric R. Application of artificial neural networks and principal component analysis on vibration signals for automated fault classification of roller element bearings. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18 (2): 299–306, http://dx.doi.org/10.17531/ein.2016.2.19.
23. Zuber N, Bajric R., Šostakov R. Gearbox faults identification using vibration signal analysis and artificial intelligence methods. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (1): 61–65.

Javier CASTILLA-GUTIÉRREZ
Juan Carlos FORTES
Department of Mining, Mechanical, Energy and Construction Engineering
Higher Technical School of Engineering, University of Huelva, Huelva, Spain

Inmaculada PULIDO-CALVO
Department of Agroforestry Sciences, Higher Technical School of Engineering
University of Huelva, Huelva, Spain

E-mails: javier.castilla@dimme.uhu.es, jcfortes@dimme.uhu.es, ipulido@uhu.es

Article citation info:
KOSOBUDZKI M, STAŃCO M. Problems in assessing the durability of a selected vehicle component based on the accelerated proving ground test. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 592–598, http://dx.doi.org/10.17531/ein.2019.4.8.

Mariusz KOSOBUDZKI
Mariusz STAŃCO

PROBLEMS IN ASSESSING THE DURABILITY OF A SELECTED VEHICLE COMPONENT BASED ON THE ACCELERATED PROVING GROUND TEST

PROBLEMY OCENY TRWAŁOŚCI WYBRANEGO ELEMENTU POJAZDU NA PODSTAWIE PRZYSPIESZONEGO TESTU PRZEBIEGOWEGO*

The paper presents the results of the analysis of the durability of elastic elements occurring in the special-purpose 4x4 off-road truck suspension using data obtained during an accelerated proving ground test conducted during off-road driving. The limitations in access to material data present at the stage of the initial selection of the component (lack of fatigue strength data) are indicated and an alternative analytical method for fatigue strength estimation is given. The differences in the obtained results and their most important sources are pointed out. A method for using a generalized durability index d as a parameter independent of the subassembly material data is also described. The indicator can be used to assess the influence of resultant loads (recorded) appearing during the vehicle operation in the determined road conditions on the durability of the subassembly under study and to associate their value with the type of the test road section.

***Keywords**: spring, stabilizer, proving ground test, accelerated tests, durability, generalized durability index, off-road truck.*

W artykule przedstawiono wyniki analizy trwałości elementów sprężystych występujących w zawieszeniu specjalnego terenowego pojazdu ciężarowego 4x4 wykorzystując dane uzyskane podczas przyspieszonego testu drogowego przeprowadzonego podczas jazdy off-road. Wskazano na występujące ograniczenia w dostępie do danych materiałowych jakie są obecne na etapie wstępnego doboru podzespołu (brak danych wytrzymałości zmęczeniowej) oraz podano alternatywną analityczną metodę szacowania wytrzymałości zmęczeniowej. Wskazano na powstające różnice w uzyskanych wynikach oraz na najważniejsze ich źródła. Przedstawiono również sposób wykorzystania uogólnionego wskaźnika trwałości d jako parametru niezależnego od danych materiałowych podzespołu, który można wykorzystać do oceny wpływu obciążeń wynikowych (rejestrowanych) powstających podczas ruchu pojazdu w ustalonych warunkach drogowych na trwałość analizowanego podzespołu i powiązać ich wartość z rodzajem testowego odcinka drogowego.

***Słowa kluczowe**: resor, stabilizator, test przebiegowy, badania przyspieszone, trwałość, uogólniony wskaźnik trwałości, ciężarowy samochód terenowy.*

1. Introduction

Vehicle durability evaluation is a very complex and challenging issue, and at the same time necessary in the process of achieving the series-production readiness of a vehicle structure [15]. In the case of complex objects, e.g., special-purpose off-road trucks that are required to be highly reliable and durable, the design and construction process is organized according to the appropriate management model. An example of such a model can be the V-model [12] developed by NASA. This model assumes that the transition to the next stage of the design and construction process is possible only when the previous step is rated positively. To have it evaluated, this is necessary to conduct appropriate tests, whose complexity and labor intensity depends on the determination of the impact degree of a given stage on the quality of the final product. Such an analysis can be performed using, for example, the Design for Six Sigma method [17], which makes it possible to indicate the accuracy that is necessary to assess individual stages of the design and construction process for the final product to have the required durability or reliability. Therefore, it is crucial to select and carry out appropriate tests that reflect with sufficient accuracy the influence of the loads predicted for the planned operating conditions on the product durability [11, 12, 14].

In the case of special-purpose off-road trucks, the selection of appropriate tests seems to be particularly tricky. These are vehicles produced in small series, designed to be driven in changing road conditions with variable load over a long period of operation (up to 30 years). It appears therefore necessary and essential to adopt several simplifying assumptions regarding, among other things, the location and conditions for testing.

Some vehicle manufacturers conduct their tests on parametrized road measurement sections, e.g., Tatra [22], which should be representative of the actual road conditions, where the degree of influence of the road profile and the vehicle traffic parameters on the value of the resulting loads and, ultimately, on the durability of the analyzed subassemblies are determined. Acceptance tests of a vehicle in running order are carried out by a designated certification body, while implementing an established testing program, on behalf of a future user. The examination results in the issuance or refusal to issue a certificate of conformity of the product with the requirements of the recipient. However, the results of these tests are available only when the vehicle is ready for production

There are also some vehicle manufacturers who do not have access to road testing centers. Thus, it is problematic to conduct tests.

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

In such a case, they are performed on selected available road sections, including public roads. However, there is a problem of correlating the loads assumed as representative (occurring in test sections of the certification body) with those used by the manufacturer. Hence, vehicle makers are looking for different parameters that can be used to compare the test conditions of the certification body with their test conditions.

Due to the limited time and financial resources, but also for example because of data shortages, testing generally leads to the ultimate limit state of the subassembly under examination or to the moment when, based on the data collected, the relationship between the conditions of use and the sustainability of the component can be established. Different models of degradation processes are used to link the resulting loads (traffic conditions) and the life span of the element. They also include those whose use does not require the knowledge of detailed material data obtained through experimental bench studies, which are very time-consuming and pricey.

The obtained results of the durability of the tested subassemblies often refer to the values describing the utilized labor resource, e.g., in units of vehicle mileage (km), engine operating hours (EOH), and others according to the future user's requirements. On the adoption of simplifying assumptions that, e.g., road test sections and established traffic parameters are constant, the acquired outcomes, enable linking the unit mileage of a vehicle with the degree of its degradation. As a result, data are received which allow for the comparison of the influence of selected types of road test sections (those of the certification body with own sections) and established traffic parameters of the vehicle on the degree of degradation of a selected component. Examples of estimating the durability of vehicle subassemblies can be found in the literature, among others in [2, 5, 10, 20]. The problem remains, however, the identification of a parameter, the determination of which could be used as a comparative indicator for the initial estimation of the component durability in connection with the selected road test section.

2. The aim and scope of the research

The research aimed to estimate the durability of selected components of an off-road truck under specified traffic conditions and to check whether it is possible to apply a generalized durability index to the initial assessment of the suitability of these components for the vehicle. The setting of the indicator mentioned above does not require full knowledge of the strength of the material of which the elements were made, which is a typical problem occurring during accelerated mileage tests. Choosing a generalized durability index and determining its value in acceptance test conditions of a certification body would allow the similar test program to be determined based on the road sections available to the manufacturer. The detailed characteristics of the generalized durability index used in the tests are not presented in this paper but are described in the publication [6].

The subject of accelerated mileage tests were elastic elements (parabolic springs, stabilizers) occurring in the suspension of a special-purpose 4x4 off-road truck. The testing was carried out under off-road traffic conditions with limited data concerning the strength of the material of which these elements were made. The vehicle manufacturer specified the limitation of the test to one type of road section.

Parabolic springs, which allow for relative movement of wheels and body in the vertical axis, and at the same time remove the freedom of movement in the other axes, and stabilizers, which reduce the lateral tilt of the body, thus improving the stability of the vehicle motion, proved to be the susceptible elements in the suspension of the analyzed vehicle [16]. The subassemblies operate in a complex stress state, but in order to simplify the tests, it is often assumed that the springs are subjected to bending and the stabilizer rods are twisted [1]. Figures 1 and 2 show the stiffness characteristics of the springs and the deflection ranges at different vehicle loads.

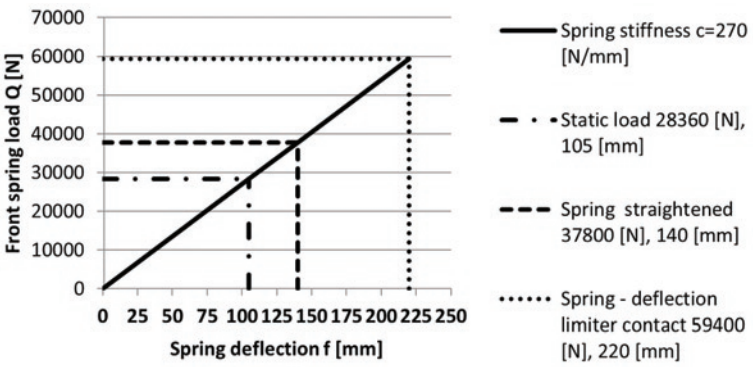


Fig. 1. Stiffness characteristics of the front axle parabolic spring

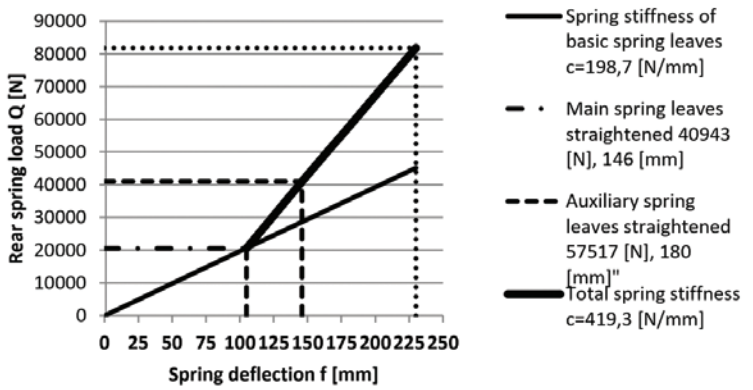


Fig. 2. Stiffness characteristics of the rear axle parabolic spring

Table 1. Summary of basic characteristic dimensions of stabilizers

	Front axle stabilizer	Rear axle stabilizer
Length of the element subject to torsion [mm]	730	820
Arm of torsional force [mm]	520	340
Diameter of the element subject to torsion [mm]	40	50
Torsional strength index of the cross-section [cm3]	6,28	12,27

The stabilizers were made of rods with a circular cross-section. The basic characteristic dimensions are shown in Table 1.

The material used in the production of these components was 51CRV4 steel (Rm=1350 MPa). The manufacturer's declaration in-

Table 2. Summary of number of samples, load levels and percentage of repetitions

Number of samples n_s	Load levels SL	Percentage of repetitions LP
12	2	83,3
12	3	75,0
24	3	87,5
24	4	83,3
24	5	79,2
24	6	75,0

dicated that the spring leaves were heat treated, and according to the standard [15], the core hardness should be in the range between 363 and 460 HB. In addition, the spring feathers on the stretched side were shot peened. With such a procedure, normal compressive stresses, which significantly reduce the values of tensile stresses arising during the component's operation, were introduced on this surface [3, 18]. Due to the lack of data on the values of those stresses as well as the depth of their introduction into the material structure, the available data, presented in, e.g., [9, 13, 15], were used to estimate them. Based on the data, it was assumed that in the unloaded state, compressive stresses might reach the value from 300 to 400 MPa, and the depth of the introduced strains can be 15 - 25 μm.

3. Fatigue strength model of the analyzed subassemblies

Testing for fatigue-limited durability of the subassemblies is a complex and time-consuming task. The sample size for experimental tests depends on the stage of the design and construction process, the number of analyzed load levels and test repetitions. In the initial step of selecting a component, the number of samples from 6 to 12 is usually enough and increases to 24 for reliability tests [1]. The number of test repetitions can be determined from the following dependency [10]:

$$LP = 100\left(1 - \frac{SL}{n_s}\right) \tag{1}$$

where: LP - percentage of repetitions, SL - number of load levels, ns - number of samples.

The percentage of repetitions in the pre-test phase is between 17 and 33. Table 2 summarizes the number of samples, load levels and repetitions for 12 and 24 samples respectively.

The presented data illustrate the time-consuming experimental tests of a subassembly performed to determine its fatigue strength characteristics. The conducted studies, which were preliminary tests of ready-made components checked by the manufacturer, were restricted to assessing the correctness of their selection for the vehicle. The tests were limited to one truck. Due to the lack of detailed data concerning fatigue strength (the experimentally determined S-N curve), it was necessary to determine the curve through theoretical calculations and to link the obtained results with the parameter connecting the component durability with the type of road test section [6].

The dependencies, which allowed to determine the fatigue graph based on a limited set of data, were used to calculate the fatigue strength of the spring. It was assumed that the determination of fatigue strength in the high cycle range, i.e., within the range from 10^3 to 10^6 cycles, was of crucial importance. The method for the determination of individual values was taken from available publications, among others [4,10].

The fatigue strength for 10^3 number of cycles was determined using the following relationships:

$$A_{NC,R} = A_{1000} \cdot C_R \tag{2}$$

where: $A_{NC,R}$ – stress amplitude for low cycle loads including reliability factor C_R, A_{1000} –stress amplitude for low cycle loads, C_R – reliability factor.

The value A_{1000} can be determined from the relationship:

$$A_{1000} = \alpha_{NC} \cdot R_m \tag{3}$$

where: R_m – limit of material strength determined in the static tensile test, α_{NC} - load type dependent coefficient for 10^3 cycles; 0.9 for bending, 0.72 for torsion.

The value of the reliability factor C_R depends on the expected operating reliability of the component. In the tests it was initially assumed that C_R=1.

The fatigue strength for 10^6 cycles was derived from the relationship in which corrective factors were taken into consideration:

$$A_{WC,R} = A_{WC} \cdot C_L \cdot C_S \cdot C_D \cdot C_R \tag{4}$$

where: $A_{WC,R}$ – stress amplitude for high cycle loads including the reliability factor C_R, A_{WC} –stress amplitude for high cycle loads, C_L – load type factor, C_S – surface condition factor, C_D – size dependent coefficient, C_R – reliability factor.

The value A_{WC} can be determined from the dependencies:

$$A_{WC} = \alpha_{WC} \cdot R_m \tag{5}$$

where: R_m – limit of material strength determined in the static tensile test, α_{WC} – coefficient depending on the type of material for 10^6 cycles; for steel (R_m <1400 MPa) it is 0.5.

The value of load type factor C_L was assumed according to data available in literature [10]. For bending C_L=1 and for twisting C_L=0,58.

The value of the surface condition factor C_S can be determined from the surface roughness measurement and the material strength value R_m. The components supplied by the manufacturer were factory protected with protective paint against the harmful effects of weather conditions. The measurement of the actual surface roughness would require the effective removal of this layer. Because of the existing limitations, the roughness was not measured and the available literature data [10] were used to determine the factor C_S. The springs were rolled and shot peened, and in this way compressive stresses were introduced into the structure of the material, thereby partially compensating the tensile stresses arising during the operation of the subassembly. The value of factor C_S equal to 0.76 was used in the calculations.

The value of the coefficient depending on the size of the C_D element was calculated from the following dependencies [10]:

$$C_D = 1{,}189 \cdot d^{-0{,}097} \tag{6}$$

where: d – element diameter, mm.

For a rectangular section element (a leaf spring), the equivalent diameter can be derived from [10]:

$$d_z = \sqrt{0{,}65 \cdot s \cdot w} \tag{7}$$

where: s - section width, w - section height.

The calculated C_D values is shown in Table 3.

Figure 3 presents the diagrams of fatigue strength of springs and stabilizers were prepared based on the determined data, which is presented in Figure 3. The determined fatigue strength values of the front and rear stabilizers are comparable, and the difference occurring in the area of unlimited fatigue strength is slight and amounts to 5 MPa. The strength values identified for this area are 247 MPa for the front stabilizer and 242 MPa for the rear stabilizer.

Table 3. The C_D values determined for the analyzed elements

	Front spring	Rear spring	Front stabilizer	Rear stabilizer
C_D	0,84	0,84	0,83	0,81

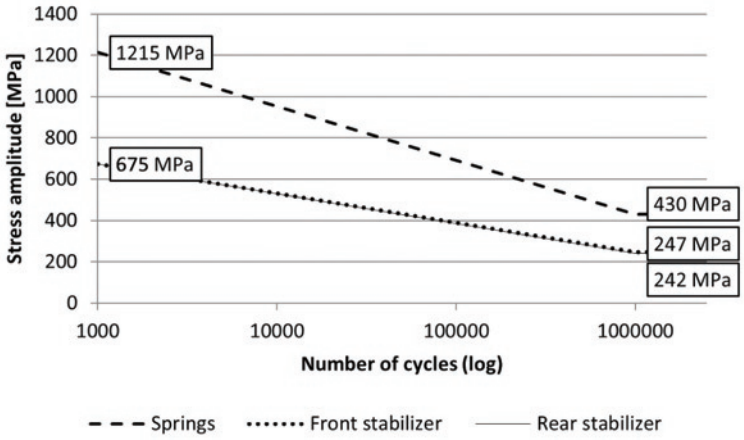


Fig. 3. Determined diagrams of fatigue strength of springs and stabilizers

The obtained diagrams of fatigue strength of the subassemblies were used to analyze the durability of the tested components.

4. The course of the tests

The examination was carried out in the training ground conditions at the University of Land Forces in Wrocław. The selected sandy off-road section was an approximately 1 km long measuring loop. Due to the nature of the unevenness, the average driving speed was about 7 km/h. It was determined from previous trips and conclusions from prior studies [7,8]. The selected road measurement section corresponded to the testing ground conditions, which are taken into consideration when designing the vehicle to the expected traffic conditions described in the vehicle exploitation profile [11]. However, that section was not parameterized. A test driver of the manufacturer drove the vehicle. The test vehicle was loaded evenly, using the total payload.

Motor vehicle springs operate in a complex stress state [1,18]. However, in accelerated mileage tests, it is difficult to record all the occurring loads and assess their influence on the fatigue life of a spring. Therefore, it is assumed that the dominant load is bending, which causes normal stresses in the cross-sections of spring leaves. In the case of stabilizers, they are designed to be torqued. Adoption of the simplifications presented causes the collection of data necessary for further analysis to be reduced to the recording of emerging stresses caused by bending of springs and torsion of stabilizers. The data reduction achieved in this way is a thoughtful step resulting from the economics of time and available resources as well as limited data on the analyzed components. Table 4 presents a set of characteristics that were available at the stage of the initial selection of subassemblies.

The measuring system used in accelerated tests of elastic vehicle components consisted of strain gauge sensors glued to the prepared surfaces of spring leaves and stabilizers (Fig. 4÷5). The strain train gauges were glued in places where the highest stress values were expected to be obtained (around the yoke fixing the spring leaves, and in the case of a stabilizer in the middle of the section subject to torsion). The choice of locations was additionally confirmed based on the FEM model of springs [19, 20], which is not a standard step.

During road tests, load courses were recorded and then filtered through Rainflow to specify load cycles. Figure 6 shows an exemplary load course of a rear axle leaf spring. Rainflow filtration was performed by determining and counting the load cycles from the recorded load course. The method is now widely used and standardized. When mounted on a vehicle, the springs are initially loaded with the vehicle's weight and freight, which affects the asymmetry of the loads generated when bending and unbending these elements while driving (shifting the mean value). The Goodman model [21] was used to take this effect into account. The Palmgren-Miner hypothesis, which assumes the linear accumulation of damages up to the limit value considered as 1, was harnessed to sum up fatigue damages. It is a model commonly used in fatigue calculations.

Fig. 5. Stabilizer with glued-on strain gauge

5. The analysis of the results obtained

The durability of the tested subassemblies was estimated from recorded mileage and theoretically determined fatigue strength, and given in units of vehicle mileage. Under the assumption that the loads occurring during the tests are representative for future predicted operating conditions, the results obtained are preliminary information used to assess the appropriateness of the choice of components for the vehicle. A significant scattering of the received values to the individual subassemblies can be observed in the summary of the results collected in Table 5. The reason for this scattering is the lack of accurate data on the actual value of pre-stresses introduced into the spring leaves, which had to be estimated.

The data presented in Table 5 indicate that the calculated spring durability is strongly influenced by the correctly assumed value of compressive pre-stress, which can be identified based on e.g., the as-

Table 4. Basic data on the analyzed subassemblies

Springs	stiffness characteristics, dimensions, weight, material, type of heat treatment and plastic processing, declared hardness on the surface, declared durability,
Stabilizers	stiffness characteristics, dimensions, weight, material, type of heat treatment and plastic processing,

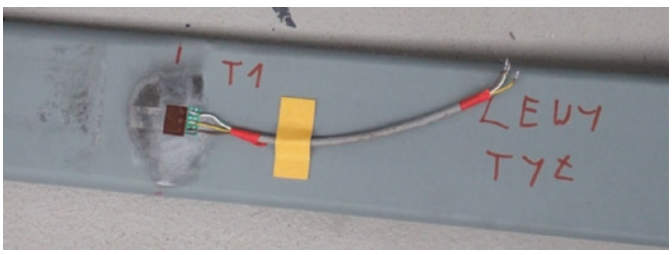


Fig. 4. Leaf spring with glued-on strain gauge

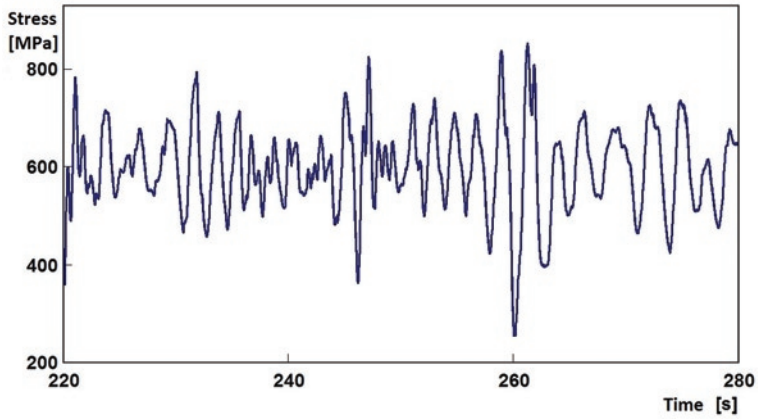


Fig. 6. Example of the stress pattern of a leaf spring installed on the rear axle of a vehicle (values do not include preliminary stresses due to peening)

sessment of the depth of changes in the microstructure of the material, which is the result of shot peening. Such an evaluation may be carried out by, among others, performing material destructive tests of a component [9, 13]. The general information provided by the manufacturer about the plastic processing, without detailed data, is insufficient for correct calculation of the component durability.

Table 6 shows how the calculated durability of the components is affected by the reduction of loads directly or indirectly influenced by the driver's driving style. From the data provided it is clear that a 5% load reduction (e.g., speed reduction, rerouting, tire pressure adjustment, etc.) can extend the life cycle of a component by approximately 50% and a 10% load reduction can increase it by ca. 100%.

The data presented show that the attempt to determine the component durability limited by fatigue strength in an accelerated mileage test poses many difficulties and may be subject to material error, e.g., due to the adoption of approximate intermediate volumes. Significant limitations in establishing the exact values include the lack of data concerning the experimentally determined fatigue strength of the subassembly, which requires approximate theoretical calculations to be made, the lack of detailed material data of the component (real value Rm, value of introduced compressive stresses and their depth) and parameters describing the condition of the top layer (roughness). Moreover, in preliminary mileage tests, when there is no access to parameterized test tracks, there is a need to compare the effects of the application of new structural solutions of subassemblies in relation to those previously used and to evaluate their work in connection with the type of road test section used by the certification body. A useful parameter in solving this type of problem may be the quantity called a generalized durability index *d* [6], which expresses numerically the overall impact of parameters describing the vehicle motion (e.g., speed, type of test section) on the durability of the component, but without reference to the material characteristics of the element.

Table 5. Summary of predicted durability of elements for different pre-stress values

Volume	Spring LP	Spring PP	Spring LT	Spring PT	Front stabilizer	Rear stabilizer
Range (excluding compressive pre- stresses) [km]	44	35	12	10	555	600
Range (initial compressive stresses 300 MPa) [km]	56698	45708	13878	11338	-	-
Range (initial compressive stresses 350 MPa) [km]	121229	100819	30231	24357	-	-
Range (initial compressive stresses 400 MPa) [km]	273596	236196	71713	56939	-	-

Table 6. Effects of load values on component durability

Subassembly	Durability at the registered load (without taking compressive pre-stresses into account)	5% reduced load durability	10% reduced load durability
Right front spring	35	48	68
Right rear spring	10	15	21
Front stabilizer	600	776	1016
Rear stabilizer	555	718	941

Table 7. Summary of the generalized durability index values for the vehicle's front and rear springs on the left and right, respectively

Value of the generalized durability index *d*	Left front spring	Right front spring	Left rear spring	Right rear spring
$d_{100\%}$	$6{,}74*10^{16}$	$8{,}48*10^{16}$	$2{,}42*10^{17}$	$2{,}80*10^{17}$
$d_{95\%}$	$4{,}71*10^{16}$	$6{,}08*10^{16}$	$1{,}72*10^{17}$	$1{,}99*10^{17}$
$d_{90\%}$	$3{,}33*10^{16}$	$4{,}30*10^{16}$	$1{,}18*10^{17}$	$1{,}40*10^{17}$
d_{300MPa}	$5{,}19*10^{13}$	$6{,}43*10^{13}$	$2{,}12*10^{14}$	$2{,}59*10^{14}$
d_{350MPa}	$2{,}43*10^{13}$	$2{,}92*10^{13}$	$9{,}73*10^{13}$	$1{,}21*10^{14}$
d_{400MPa}	$1{,}07*10^{13}$	$1{,}24*10^{13}$	$4{,}10*10^{13}$	$5{,}19*10^{13}$

The concept of using a generalized durability index d is described in [6] and is based on the determination of the value of the expression:

$$d = \sum n_i A_i^{\beta} \tag{8}$$

where: d - generalized durability index (pseudo damage), A_i - load amplitude determined by, e.g., Rainflow method, n_i - number of load cycles with A_i amplitude, β - fatigue curve slope coefficient (it may be assumed that for elements performed without special finishing operations (e.g., grinding, polishing) the coefficient β=5).

The described generalized durability index d was used to present the differences in the loads of the same components on the left and right respectively. Examples of the calculation results are given in Table 7.

The values of the generalized durability index d presented in Table 7 apply to cases where the values of measured stresses (d100%), stresses reduced by 5% and 10% (d95%, d90%) and initial compressive stresses (300 MPa, 350 MPa, and 400 MPa, respectively) were considered. The increasing value of the parameter d indicates a more destructive course of loads. The data presented in Table 7 show that the front right spring, which is the same as the front left one, was subjected to more destructive loads during the tests. Similarly, the right rear spring was more fatigue loaded than the rear left one. One can also see that the elastic components in the front axle suspension are more durable than those in the rear axle. However, the received values for the generalized durability index d of the subassembly do not represent the actual life cycle of the component, but only constitute a numerical representation (easy to compare) of whether the loads acting on the element are more or less destructive under given traffic conditions compared to another component of the same type.

6. Conclusion

The primary objective of the research was to identify the loads acting on the spring components of the suspension and to estimate their durability limited by fatigue strength, as shown in Table 5 and used as preliminary data to check the suitability of these subassemblies for the vehicle. An additional aim was to indicate a parameter, the use of which would allow for the assessment of the extent to which the traffic conditions and the type of road measuring section influence the value of loads on the selected elements, thus limiting their durability.

The durability of the analyzed components is crucial for the estimation of vehicle reliability, which is understood as a technical system whose loads resulting from traffic conditions vary widely (from driving on hard-surfaced roads with no cargo to off-road driving with freight). The analyses presented were based on limited data available at the stage of the initial selection of a new subassembly for the vehicle. The obtained results of component durability are presented concerning the theoretical driving range of the car, which is an effective comparison parameter. Due to limited data, different load values (as the result of possible changes in the driver's driving style) and initial compression stresses of spring leaves were used for calculations, thereby showing how they affect the vehicle mileage being analyzed.

The tests were limited to only one vehicle (one set of analyzed subassemblies) moving at a set speed in selected road conditions. Therefore, the results obtained are only a preliminary material for further analysis. However, it is worth noting that the use of the proposed generalized durability index d makes the initial comparison of the durability of individual vehicle springs possible. The distinction of the degradation degree of the same springs, but differently loaded (which stems from the non-identical shape of the ground under each wheel during the journey) indicates that the values of parameter d determined for the same component (spring) in various road conditions (test sections) can also be compared. If the parameter d is additionally normalized and its value is reduced to the unit length of the measurement distance (e.g., to 1 km), it will be possible to estimate the degradation degree of the same component in different traffic conditions and on varying test sections. This gives reason to believe that it is possible to reproduce the effect of the loads recorded on one test section (e.g., of a certification body) with another available test section (available from the vehicle manufacturer), which would be an innovative use of the parameter d identified on the basis of the transformed Basquin equation. Confirmation of this assumption will, however, require additional testing.

References

1. ASTM E739-10 - Standard practice for statistical analysis of linear or linearized stress-life (S-N) and strain-life (ε-N) fatigue data. USA: ASTM International, West Conshohocken, PA, 2015.
2. Johannesson P, Speckert M. (editors) Guide to load analysis fo r durability in vehicle engineering. London: John Willey & Sons, 2014, https://doi.org/10.1002/9781118700518.
3. Hryciów Z, Krasoń W, Wysocki J. The experimental tests of the friction coefficient between the leaves of the multi-leaf spring considering a condition of the friction surfaces. Eksploatacja i Niezawodność - Maintenance and Reliability 2018; 20(4): 682-688, https://doi.org/10.17531/ein.2018.4.19.
4. Kocańda S, Szala J. Podstawy obliczeń zmęczeniowych. Warszawa: PWN, 1991.
5. Kosobudzki M. Metoda szacowania trwałości ustroju nośnego pojazdu wysokiej mobilności. Rozprawa doktorska. Politechnika Wrocławska, 2013.
6. Kosobudzki M, Smolnicki T. Generalized vehicle durability index for different traffic conditions, AIP Conference Proceedings 2019; 2078: 020017(1-6), https://doi.org/10.1063/1.5092020.
7. Kosobudzki M, Stanco M. The experimental identyfication of torsional angle on a load-carrying truck frame during static and dynamic tests. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(2): 285-290, https://doi.org/10.17531/ein.2016.2.17.
8. Kosobudzki M. The use of acceleration signal in modeling process of loading an element of underframe of high mobility wheeled vehicle. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(4): 595-599.
9. Kukiełka L, Bartosik P, Szyc M. Optymalizacja procesu kulowania strumieniowego ze względu na naprężenia wynikowe, Archiwum Technologii Maszyn i Automatyzacji, KBM PAN_Oddział w Poznaniu 2010; 30/1: 117-126.
10. Lee Y-L, Pan J, Hathaway R, Barkey M. Fatigue testing and analysis. Theory and practice. Elsevier, 2005.
11. MIL-STD_810G - Environmental engineering considerations and laboratory tests. USA: Department of Defense test method standard, 2008.
12. NASA System Engineering Handbook. USA: NASA Headquarters, Washington, 2007.
13. Nasiłowska B, Bogdanowicz Z, Brzeziński M, Mońka G, Zasada D. Wpływ kulowania na strukturę, mikrotwardość i naprężenia własne stali

austenitycznej 1.4539. Biuletyn WAT 2015; 64/2: 103-110, https://doi.org/10.5604/12345865.1157224.
14. Norma Obronna NO-06-A101 - Uzbrojenie i sprzęt wojskowy. Ogólne wymagania techniczne, metody kontroli i badań. Postanowienia ogólne
15. Norma PN-90/S-47250 - Pojazdy samochodowe i przyczepy. Resory piórowe. Wymagania i badania.
16. Reimpell J, Betzler J. Podwozia samochodów. Podstawy konstrukcji. Warszawa: Wydawnictwo Komunikacji i Łączności WKŁ, 2008.
17. Rusiński E, Koziołek S, Jamroziak K. Quality assurances metod for the design and manufacturing process of armoured vehicles. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2009; 43(3): 70-77.
18. Spring design manual. AE-21. USA: SAE International, 1996.
19. Stańco M, Iluk A. Numeryczno - doświadczalna analiza wytężenia resoru parabolicznego pojazdu ciężarowego. Materiały konferencyjne XVI Konferencji Naukowo - Technicznej TKI2016 - Techniki Komputerowe w Inżynierii, 18-21.10.2016.
20. Stańco M. Analysis of the influence of leaf geometry on stiffness and effort of the heavy-duty spring. In: Rusiński E, Pietrusiak D. (editors) Proceedings of the 14th International Scientific Conference - Computer Aided Engineering. Springer International Publishing, 2018, https://doi.org/10.1007/978-3-030-04975-1_84.
21. Łagoda T, Macha E. Trwałość zmęczeniowa maszyn - laboratorium. Opole: Politechnika Opolska, 2005.
22. www.tatratrucks.com/your-tatra-partner/tatra-testing-grounds/

Mariusz KOSOBUDZKI
Mariusz STAŃCO
Department of Mechanics, Materials Science and Engineering,
Wroclaw University of Science and Technology
ul. Łukasiewicza 7/9, 50-371 Wrocław

E-mails: mariusz.kosobudzki@pwr.edu.pl, mariusz.stanco@pwr.edu.pl

Article citation info:
YOU L, ZHANG J, LI Q, YE N. Structural reliability analysis based on fuzzy random uncertainty. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 599–609, http://dx.doi.org/10.17531/ein.2019.4.9.

Lingfei YOU
Jianguo ZHANG
Qiao LI
Nan YE

STRUCTURAL RELIABILITY ANALYSIS BASED ON FUZZY RANDOM UNCERTAINTY

ANALIZA NIEZAWODNOŚCI STRUKTURALNEJ W OPARCIU O ROZMYTĄ NIEPEWNOŚĆ LOSOWĄ

To address the fuzzy random uncertainty in structural reliability analysis, a novel method for obtaining the membership function of fuzzy reliability is proposed on the two orders four central moments (TOFM) method based on envelope distribution. At each cut level, the envelope distribution is first constructed, which is a new expression to describe the bound of the fuzzy random variable distribution. The central moments of the bound distribution are determined by generating samples from the envelope distribution, and they are used to calculate the central moments of the limit state function based on the first two orders of the Taylor expansion. Thereafter, the modern approximation method is used to approximate the polynomial expression for the limit state function probability density function (PDF) by considering the central moments as constraint conditions. Thus, the reliability boundaries can be calculated under the considered cut level, and the membership function of the fuzzy reliability is subsequently obtained. Three examples are evaluated to demonstrate the efficiency and accuracy of the proposed method. Moreover, a comparison is made between the proposed method, Monte Carlo simulation (MCS) method, and fuzzy first-order reliability method (FFORM). The results show the superiority of the proposed method, which is feasible for the analysis of structural reliability with fuzzy randomness.

Keywords: *fuzzy random uncertainty, approximation method, envelope distribution, structure, cut level.*

W pracy przedstawiono metodę, która pozwala na uwzględnienie rozmytej niepewności losowej w strukturalnej analizie niezawodności. Zaproponowana metoda określania funkcji przynależności niezawodności rozmytej wykorzystuje cztery momenty centralne dwóch rzędów czy czwarte momenty centralne drugiego rzędu obliczane w oparciu o rozkład obwiedni. Dla każdego poziomu cięcia, najpierw konstruuje się rozkład prawdopodobieństwa obwiedni, za pomocą którego opisuje się granice rozkładu rozmytych zmiennych losowych, a momenty centralne rozkładu ograniczonego wyznacza się poprzez generowanie prób z rozkładu obwiedni. Następnie, stosując nowoczesną metodę optymalnej aproksymacji, otrzymuje się aproksymowane wyrażenie wielomianowe funkcji gęstości prawdopodobieństwa rozkładu obwiedni, gdzie momenty centralne stanowią warunki ograniczające, które pozwalają aproksymować niezawodność za pomocą rozwinięcia Taylora drugiego rzędu funkcji stanu granicznego. W ten sposób granice niezawodności oblicza się na rozważanym poziomie cięcia, a następnie otrzymuje się funkcję przynależności niezawodności rozmytej. W artykule przeanalizowano trzy przykłady, na podstawie których wykazano skuteczność i trafność proponowanej metody. Przeprowadzono także porównanie z metodą symulacji Monte Carlo oraz metodą analizy rozmytej niezawodności pierwszego rzędu. Wyniki wskazują na wyższość omawianej metody, która pozwala analizować niezawodność strukturalną w warunkach losowości rozmytej.

Słowa kluczowe: *rozmyta niepewność losowa, metoda aproksymacji, rozkład obwiedni, struktura, poziom cięcia.*

1. Introduction

There are several uncertainties with respect to the analysis of structural reliability, and the fluctuations due to the uncertainty have a significant influence on the performance of structure products, which increases the requirements of the uncertainty analysis method for achieving reliable structures.

Traditionally, uncertainty is classified into two major categories, namely, aleatory or epistemic. Aleatory uncertainties in reliability analysis have been successfully addressed using the probability theory, which requires completely statistical information based on probability distributions to describe the aleatory uncertainties [2,5,30]. The probabilistic reliability analysis methods with random variables include the moments method [3, 10, 40], response surface method [16, 17], Monte Carlo method [36], and direct integration method [38]. Although the probabilistic methods have been successfully applied, the quality of the input information should be statistically guaranteed by a sufficiently large set of sample elements to verify the used distributions.

In contrast to aleatory uncertainties, epistemic uncertainties are knowledge-based and arise from imprecise modelling, simplification, and limited data availability [11]. There are several approaches for modelling epistemic uncertainties, such as the convex model method [9], possibility theory method [18], interval modelling [15, 27], evidence theory [1, 37], and uncertainty theory [20]. As their representative, the fuzzy sets theory is widely used for reliability analysis [7, 32, 39]. By the membership functions [28, 31], fuzzy reliability analysis can account for inaccuracies and uncertainty in data, which typically occurs when insufficient data is available to provide a useful statistical description.

However, with significant research on physical modeling and reliability analysis, it is found that aleatory and epistemic uncertainties do not exist alone, i.e., certain information, precise values, and completely obscure information do not exist. Thus, the concept of the

fuzzy random variable was proposed [14], where uncertain structural parameters governed by probability distributions with fuzzy parameters were introduced. Moreover, the fuzzy random variable reconciles aleatory and epistemic uncertainties, allowing an uncertain expression with random distribution and incomplete information to be constructed.

Willner [34] proposed an engineering concept to address fuzzy randomness. Möller et al. [25,26] presented a method for describing and predicting fuzzy time-series based on fuzzy random uncertainties. Liu et al. [21] used fuzzy random variables as basic variables to establish a relationship between fuzzy random variables, in addition to fuzzy random events. Körner [13] evaluated the properties of the variations in fuzzy random variables, and then applied to linear regression and limit theorems. Möller et al. [24] introduced a method for estimating the membership function of the safety index under the consideration of fuzzy randomness. A fuzzy first-order reliability method (FFORM) was developed using fuzzy random variables. Terán [33] presented probabilistic results toward a framework for modelling measurements based on fuzzy random variables. Wang et al. [35] solved the time dependent reliability problem for systems with fuzzy random uncertainties using saddle point approximation simulations. Koç et al. [12] used the theory of fuzzy random variables with fuzzy Monte Carlo simulations for reliability-based risk analysis of a rubble-mound breakwater. Shapiro [29] modelled the future lifetime as a fuzzy random variable, where the essential feature of the model was combined the stochastic component of mortality with the fuzzy component. In the study conducted by Jahani et al. [8], uncertain variables were modeled as fuzzy random variables. In addition, an interval Monte Carlo simulation (IMC) and the interval finite element method were used to evaluate the failure probability. Hryniewicz [6] presented a Bayesian approach to analysis the reliability under fuzzy random data. Li et al. [19] proposed a fuzzy reliability calculation method based on the error synthesis principle for fuzzy random uncertainty inputs.

The abovementioned methods can be divided into three categories: namely, iteration algorithms, sampling algorithms, and approximation algorithms. However, the application of fuzzy random uncertainties in addressing the reliability presents several problems when the abovementioned methods are used. With the combination of an iteration algorithm and traditional reliability algorithm, the calculation efficiency is not satisfied, and the accuracy is insufficient for high nonlinearity limit state functions. Moreover, sampling algorithms requires significant operations in the membership interval, for which the efficiency is insufficiently low for complex structures. For the application of approximation algorithms, it requires cumbersome transformations, which has a tremendous possibility of improvement.

Therefore, a novel structure reliability analysis method on TOFM based on envelope distribution is developed by combining the modern approximation algorithm, which considers the basic input variables as fuzzy random variables, and reliability analysis is expressed with respect to fuzzy numbers using the α cut level approach. In this study, modern approximation algorithms such as the maximum entropy model [1] and optimal square approximation method [22,41] were used to approximate the fuzzy probability density function (FPDF) with fuzzy random variable inputs. Only the central moments are used in the approximation without considering the actual distribution. At each cut level, a new measure distribution named envelope distribution is used to establish an accurate description for the envelope of the fuzzy random distribution, which is the boundary of the distribution family of fuzzy random variables. In addition, the first four central moments of the envelope distribution are obtained using a statistical method and then the moments of limit state function are approximated based on envelope distribution moments according to the first two orders of magnitude of the Taylor expansion on limit state function. Thereafter, by considering the central moments as the constrained conditions, the undetermined polynomial coefficients are fitted by employing the modern approximation method. Hence, the approximated polynomial expression of the limit state function PDF boundary is obtained. Thereafter, the boundary of the reliability membership function is calculated, and the fuzzy reliability is obtained by the application of the abovementioned operation at each cut level.

Compared with traditional methods, the proposed method can solve the drawback of high computation loads, poor accuracy, and instability due to fuzzy random uncertainties. It facilitates reliability analysis without iterative algorithms at each cut level, whereas the classical reliability analysis method requires computationally complex searches or optimization procedures. Furthermore, the proposed method only uses moments obtained from the statistical analysis of basic data, which is convenient for practical operations.

This article is structured as follows. Section 2 presents a brief introduction to the fuzzy random variable. In Section 3, the concept of moment generation based on the sampling of the envelope distribution is presented. Section 4 presents a discussion on fuzzy reliability, in addition to modern approximation algorithms using the central moments of envelope distribution. Finally, in Section 5, three examples are provided to illustrate the method.

2. Fuzzy Random Variable and Reliability

A fuzzy random variable $\tilde{x}$ is a random variable for which its distribution parameters are fuzzy numbers. $\tilde{x}$ can be defined on a fuzzy probability space Ω, F, P, wherein Ω is the space of the fuzzy random elementary events, and F and P are the subsets and fuzzy probability measure, respectively. A fuzzy random variable $\tilde{x}$ defines a mapping relationship from $(\Omega,\mu(\Omega))$ to $\left(R^n,\mu\left(R^n\right)\right)$, i.e., $(\Omega,\mu(\Omega)) \rightarrow \left(R^n,\mu\left(R^n\right)\right)$ [35][4], where $\mu(\cdot)$ is the membership degree. Each fuzzy random variable $\tilde{x}$ contains a basic realization random variable x as the initial of $\tilde{x}$. The α cut level approach is used to conduct fuzzy arithmetic operations. Hence, the fuzzy probability cumulative distribution function (FPCDF) of a fuzzy random variable can be expressed as follows:

$$\tilde{F}(x)=\left\{\left(F_\alpha(x),\mu\left(F_\alpha(x)\right)\right)\middle|F_\alpha(x)=\left[F_\alpha^L(x),F_\alpha^U(x)\right],\mu\left(F_\alpha(x)\right)=\alpha\forall\alpha\in(0,1]\right\} \tag{1}$$

where $F_\alpha(x)$ is the FPCDF under the α cut level, and $F_\alpha^L(x)$ and $F_\alpha^U(x)$ are the lower and upper bounds of $F_\alpha(x)$, respectively. There is a set of distributions under different membership degrees. Fig. 1 presents a fuzzy random variable $\tilde{x}$ with FPCDF $\tilde{F}(x)$ and FPDF $\tilde{f}(x)$. The dashed and solid lines indicate probability functions with fuzzy parameters that correspond to membership degrees of value 0 and 1, respectively.

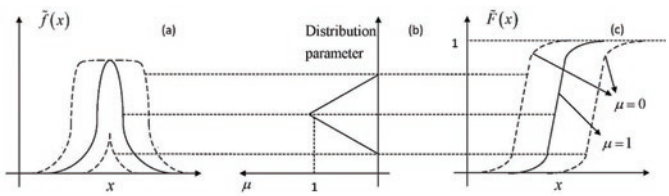


Fig. 1. (a) Fuzzy probability density function; (b) fuzzy parameter; (c) fuzzy probability distribution functions.

Fuzzy random reliability is based on the use of fuzzy random variables as the basic variables for the reliability problem, which is meas-

ured using the membership degree [34,23]. The limit state function y of the reliability model is defined as:

$$y = g\left(\tilde{\mathbf{X}}\right) = g\left(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n\right) \tag{2}$$

where $\tilde{\mathbf{X}} = \{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n\}$ are n dimensional fuzzy random variables, which have FPDFs of $f_\alpha\left(x_i\right)\left(i = 1,2,\ldots,n\right)$ at the α cut level. The fuzzy theory-based reliability defined as $\tilde{R} = \left\{\left\{\left((P_r)_\alpha, \mu\left((P_r)_\alpha\right)\right) \middle| (P_r)_\alpha = \Pr\{g\left(\mathbf{X}\right) > 0\}, \mathbf{X} \sim f_\alpha\left(x_i\right)\left(i = 1,2,\ldots,n\right)\right\}\right\}$. It represents the influence of fuzziness on reliability based on different membership levels. Once the basic variables are defined using fuzzy membership functions at various membership levels, the reliability interval at the α cut level $(P_r)_\alpha = \left[(P_r)_\alpha^L, (P_r)_\alpha^U\right]$ can be obtained with respect to $F_\alpha^L\left(x\right)$ and $F_\alpha^U\left(x\right)$, respectively; $(P_r)_\alpha^L$ and $(P_r)_\alpha^U$ are the lower and upper bounds of P_r at the α cut level, respectively.

3. Moment Generation Based on Envelope Distribution

Moment generation based on sampling from the envelope distribution is presented in this section. The envelope distribution is an envelope line that consists of the upper and lower boundaries of the FPCDFs of fuzzy random variables. The objective of envelope distribution is to comprehensively describe the boundary distribution of $\tilde{F}\left(x\right)$, and the central moments can be obtained using a statistical method. This is used for reliability analysis, which will be discussed in the next section.

For convenience, the process of generating an envelope distribution is illustrated by assuming the basic realization of a fuzzy random variable as normal distribution, as well as the other distribution. $\tilde{\mu}$ and $\tilde{\sigma}$ are the fuzzy mean value and fuzzy standard deviation of a fuzzy random variable $\tilde{x} = N(\tilde{\mu}, \tilde{\sigma})$, respectively. All the membership functions are assumed to be fuzzy triangular number. Hence, the fuzzy mean and standard deviation can be expressed as $\tilde{\mu} = \mu_{Low}; \mu_{Mid}; \mu_{Up}$ and $\tilde{\sigma} = \sigma_{Low}; \sigma_{Mid}; \sigma_{Up}$, respectively, where the subscripts Low, Mid, and Up are the lower bound, median bound, and upper bound, respectively (in the following, these labels will be written as superscripts once the cut level expression is introduced). A fuzzy random variable with $\tilde{\mu} = \langle -0.5, 0, 0.5\rangle$ and $\tilde{\sigma} = \langle 0.9, 1, 1.1\rangle$ is generated in MATLAB as an example, which is shown in Fig. 2, where the black line in Fig. 2(a) and the middle black line in Fig. 2(b) correspond to a membership degree of 1. According to the curves in Fig. 2(b), the boundary of the FPCDF is found to be an envelope of a set of curves. The upper and lower black lines in Fig. 2(c) are the envelope curves can be obtained by following operation: at each α cut level, after the bound of the interval numbers, $\left[\mu_\alpha^L, \mu_\alpha^U\right]$ and $\left[\sigma_\alpha^L, \sigma_\alpha^U\right]$ are obtained, the upper bound PDF($F_\alpha^U\left(x\right)$) of the fuzzy random variables is constructed by sampling from $x \sim N\left(\mu_\alpha^U, \sigma_\alpha^L\right)$ on the left side of μ_α^U and from $x \sim N\left(\mu_\alpha^U, \sigma_\alpha^U\right)$ on the right side of μ_α^U. The set of sampling points is defined as $\bar{\mathbf{X}}_\alpha = \{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n\}$. In contrast, the lower bound PDF ($F_\alpha^L\left(x\right)$) is constructed by sampling from $x \sim N\left(\mu_\alpha^L, \sigma_\alpha^U\right)$ on the left side of μ_α^L, and from $x \sim N\left(\mu_\alpha^L, \sigma_\alpha^L\right)$ on the right side of μ_α^L. These sampling points are defined as $\bar{\mathbf{X}}_a = \{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n\}$. It should note that the envelope curve can be directly computed from CDFs if it could be expressed expediently, but in some cases the proposed generation method is really needed:1. The expression of CDF is very complex, such as the marginal distribution under the joint distribution of polar diameter and polar angle in two-dimensional irregular walking issue.2. Those truncated distributions that are hard to express CDF, which is applied widely in engineering 3. The distribution which central moments can not be expressed, e.g., a Cauchy distribution.

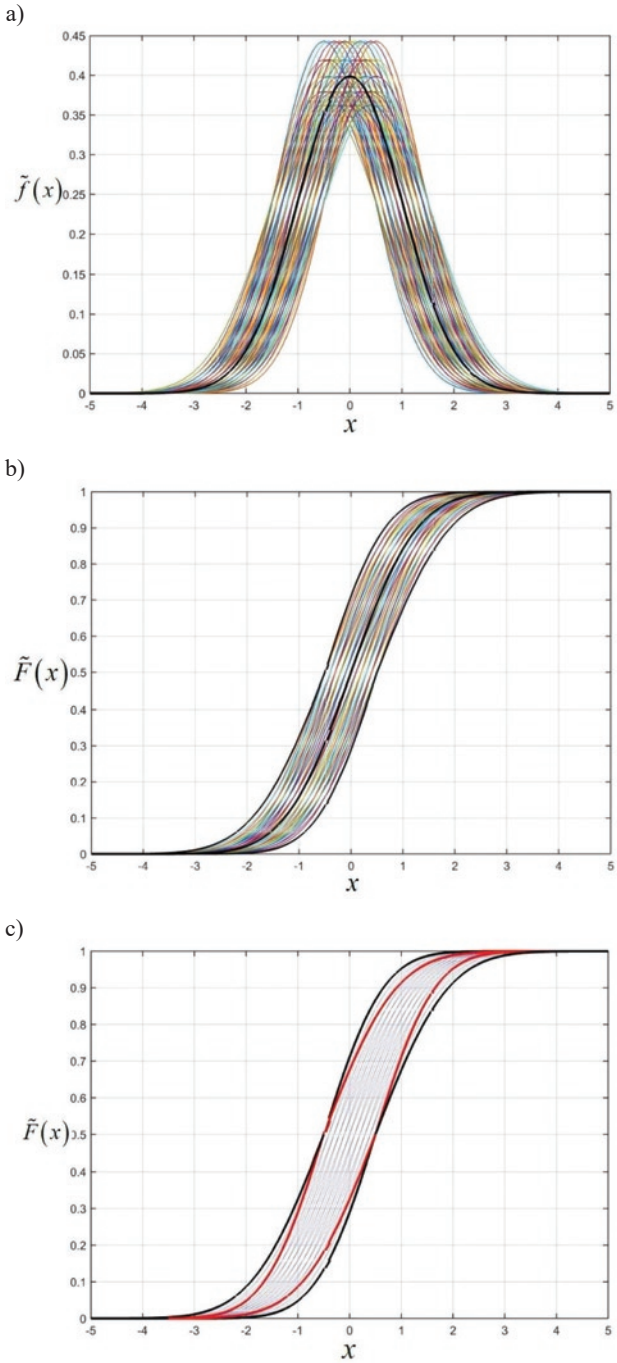


Fig. 2. (a) FPDF of fuzzy random variable; (b) FPCDF fuzzy random variable(c) envelope distribution

The envelope curve can encapsulate the boundaries of the distribution family. If only the upper and lower bound of the mean value are considered as sampling centers in the entire region, instead of separate on both sides of μ_α^U and μ_α^L, e.g., $x \sim N\left(\mu_\alpha^U, \sigma_\alpha^L\right)$ or $x \sim N\left(\mu_\alpha^U, \sigma_\alpha^U\right)$. This will produces inaccurate results, as indicated with the red line in Fig. 2(c).

The ith central moments of $F_\alpha^U\left(x\right)$ and $F_\alpha^L\left(x\right)$ are expressed as $\nu_{x_\alpha^U i}$ and $\nu_{x_\alpha^L i}$, respectively, which used in next section. Based on $\bar{\mathbf{X}}_\alpha$ and $\bar{\mathbf{X}}_a$, they can be calculated using a statistical method or a

simple method for generating central moments, e.g., a universal generating function. Thereafter, the modern approximation method is employed to calculate the reliability interval at a given cut level.

4. Modern approximation method based on central moments of the envelope distribution

In this section, modern approximation methods that considers the central moments of the envelope distribution as constraint conditions is presented, which are used to approximate the fuzzy reliability $(P_r)_\alpha = \left[(P_r)_\alpha^L, (P_r)_\alpha^U\right]$ composed of different α cut levels. TOFM based on the envelope distribution can prevent large amount of iterations and complex transformations. As typical modern approximation algorithms, the maximum entropy model and optimal square approximation methods are employed in TOFM in this study. These methods are extensively used due to their satisfactory fitting effect and easy implementation. The limit state function is defined as $Z = g\left(\tilde{\mathbf{X}}\right)$, where $\tilde{\mathbf{X}}$ is the set of fuzzy random variables $\tilde{\mathbf{X}} = \left(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n\right)$. Based on the FCDF bounds $F_\alpha^U(\mathbf{X})$ and $F_\alpha^L(\mathbf{X})$ of $\tilde{\mathbf{X}}$, the upper and lower bounds of Z_α^U and Z_α^L are approximated by the first two orders of magnitude of the Taylor expansion at the MPP(most probable point, i.e. the point of greatest contribution to failure probability) $\mathbf{x}_\alpha^{*U}$ ($\mathbf{x}_\alpha^{*L}$) as follows:

$$\begin{cases} Z_\alpha^U \approx g_{X_\alpha^U}\left(\mathbf{x}_\alpha^{*U}\right) + (\mathbf{X}_\alpha^U - \mathbf{x}_\alpha^{*U})^T \nabla g_{X_\alpha^U}\left(\mathbf{x}_\alpha^{*U}\right) + \frac{1}{2}\left(\mathbf{X}_\alpha^U - \mathbf{x}_\alpha^{*U}\right)\nabla^2 g_{X_\alpha^U}\left(\mathbf{x}_\alpha^{*U}\right)\left(\mathbf{X}_\alpha^U - \mathbf{x}_\alpha^{*U}\right) \\ Z_\alpha^L \approx g_{X_\alpha^L}\left(\mathbf{x}_\alpha^{*L}\right) + (\mathbf{X}_\alpha^L - \mathbf{x}_\alpha^{*L})^T \nabla g_{X_\alpha^L}\left(\mathbf{x}_\alpha^{*L}\right) + \frac{1}{2}\left(\mathbf{X}_\alpha^L - \mathbf{x}_\alpha^{*L}\right)\nabla^2 g_{X_\alpha^L}\left(\mathbf{x}_\alpha^{*L}\right)\left(\mathbf{X}_\alpha^L - \mathbf{x}_\alpha^{*L}\right) \end{cases} \tag{3}$$

where $\mathbf{X}_\alpha^U = \left(x_{1\alpha}^U, x_{2\alpha}^U, \ldots, x_{n\alpha}^U\right)$ and $\mathbf{X}_\alpha^L = \left(x_{1\alpha}^L, x_{2\alpha}^L, \ldots, x_{n\alpha}^L\right)$ are the variables that with distributions $F_\alpha^U(\mathbf{X})$ and $F_\alpha^L(\mathbf{X})$, respectively. $\nabla g_X(\cdot)$ is the partial derivative vector, $\nabla^2 g_X(\cdot)$ is the Hessian matrix. For convenience, the proposed method will be illustrated with the upper bound PDF approximation in the following. A standard normally distributed random variable Z_α^U can be normalized as $Y_\alpha^U = \left(Z_\alpha^U - \mu_{Z_\alpha^U}\right) / \sigma_{Z_\alpha^U}$, where $\mu_{Z_\alpha^U}$ and $\sigma_{Z_\alpha^U}$ are the mean and standard deviation of Z_α^U, respectively. Hence, the ith central moments of the upper bound at the α cut level can be calculated using $v_{Y_\alpha^U i} = E[\left(Y_\alpha^U\right)^i] = \int_{-\infty}^{+\infty} (y)^i f_{Y_\alpha^U}(y)dy$, which is a function of $v_{x_{j\alpha}^U i}, i = 0,1,2,3,4, j = 1,2,\ldots,n$ according to Eq. (3). In addition, $v_{x_{j\alpha}^U i}, i = 0,1,2,3,4, j = 1,2,\ldots,n$ are generated from $\bar{\mathbf{X}}_{j,\alpha}\, j = 1,2,\ldots,n$ using the method discussed in Section 3. Thereafter, TOFM based on the envelope distribution can be implemented using modern approximation algorithms. To clearly demonstrate the proposed algorithm, the basic theories of the maximum entropy model and optimal square approximation method are briefly reviewed and combination with central moments are investigated in the following subsection.

4.1. Maximum entropy model based on central moments

Shannon entropy is a measure of the degree of uncertainty of an event prior to its occurrence. Moreover, it is a measurement of the amount of information obtained from the event after the event (information content). Under given conditions, there is a distribution of all possible probability distributions, which maximizes the information entropy. This is referred to as the Jaynes maximum entropy principle. Under the constraint of known information, the information entropy is greatest, and the probability distribution is the least biased. The entropy of the continuous random variable x with PDF $f(x)$ is defined as [16]:

$$H = -c\int_{-\infty}^{+\infty} f_X(x) ln f_X(x) dx \tag{4}$$

where H is referred to as the Shannon entropy and c is Boltzmann's constant, which is greater than 0. Considering the central moments $v_{Z_\alpha^U i}(i = 0,1,2,3,4)$ of the limit state function Z_α^U as the constraint condition after normalization, the maximum entropy model of the upper bound of Z at the α cut level can be expressed as follows:

$$\begin{cases} \max H = -c\int_{-\infty}^{+\infty} f_\alpha^U(z)\ln f_\alpha^U(z)dz \\ s.t. \quad E[\left(Y_\alpha^U\right)^i] = v_{Y_\alpha^U i}, i = 0,1,2,3,4 \end{cases} \tag{5}$$

The Lagrange multiplier method is therefore employed to solve the maximum entropy model, i.e., $L = H + \sum_{i=0}^{4}\lambda_i\left(E[\left(Y_\alpha^U\right)^i] - v_{Y_\alpha^U i}\right)$. The undetermined constant is defined as $a_0 = 1 - \frac{\lambda_0}{c}$, where $a_i = -\frac{\lambda_0}{c}(i = 1,2,3,4)$, and the approximate expression of the probability density function of the limit state function is:

$$f_{Y_\alpha^U}(y) = \exp\left(-\sum_{j=0}^{4} a_j y^j\right) \tag{6}$$

On the other hand, the first four moments of the upper bound at the α cut level are calculated from $v_{x_{j\alpha}^U i}, i = 0,1,2,3,4, j = 1,2,\ldots,n$, which are the moments of the envelope distribution as mentioned above. Substituting Eq. (6) into Eq. (5) yields Eq. (7):

$$\int_{-\infty}^{+\infty} y^i \exp\left(-\sum_{j=0}^{4} a_j y^j\right)dy = v_{Y_\alpha^U i}, i = 0,1,2,3,4 \tag{7}$$

The polynomial fitting coefficients $a_0, a_1, \ldots, a_m$ of $f_{Y_\alpha^U}(y)$ could then be determined.

4.2. Optimal square approximation model based on central moments

The theoretical basis of the optimal square approximation method is as follows. If the central moments of two random variables are equal at each order, they have the same probability distribution characteristics and eigenvalues. The undetermined coefficients of the PDF polynomials can be obtained by considering the central moments of each order as constraints in a given inner product space, thus determining the probability distribution [22, 41].

According to the above analysis, the FPDF bound ($f_\alpha^U(z)$) must be approximated at the given cut level. The optimal square approximation model involved in fuzzy random variables can be expressed as follows:

$$\begin{cases} \min I = \int_a^b \left[p_\alpha^U(z) - f_\alpha^U(z) \right]^2 \rho(z) dz \\ s.t. \quad p_\alpha^U(z) = \sum_{i=0}^m \lambda_i p_{\alpha,i}^U(z) \end{cases} \tag{8}$$

where $f_\alpha^U(z)$ is the upper bound PDF of the limit state function at the α cut level, $p_\alpha^U(z)$ is the approximate polynomial expression of $f_\alpha^U(z)$, $p_{\alpha,i}^U(z)(i=0,1,2,\ldots,m)$ is a continuous function of $m+1$ linearly independent functions based on the limit state function interval $[a,b]$, $\lambda_i(i=0,1,\ldots,m)$ is the respective coefficient, and $\rho(x)$ is the weight function of the power on interval $[a,b]$.

The necessary condition $\frac{\partial I}{\partial \lambda_i} = 0$ for the extremum of a multivariate function can be used to determine the system of linear equations with coefficients $\lambda_0, \lambda_1, \ldots, \lambda_m$, as follows:

$$\mathbf{A}\boldsymbol{\lambda} = \boldsymbol{B} \tag{9}$$

where the respective components of the matrix elements and vectors are:

$$A_{ij} = \int_a^b p_{\alpha,i}^U(z) p_{\alpha,j}^U(z) \rho(z) dz = 0, i, j = 0,1,\ldots,m \tag{10}$$

$$B_i = \int_a^b f_\alpha^U(z) p_{\alpha,i}^U(z) \rho(z) dz = 0, i = 0,1,\ldots,m \tag{11}$$

Given that $p_{\alpha,0}^U(z), p_{\alpha,1}^U(z), \ldots, p_{\alpha,m}^U(z)$ are linearly independent, and $\mathbf{A}$ is an $m+1$ order non-singular matrix, Eq. (9) has unique solutions. Let $p_{\alpha,i}^U(z) = z^i, i = 1,2,\ldots m$ and $\rho(z) = 1$; $\mathbf{A}$ and $\mathbf{B}$ can then be calculated using Eqs. (12) and (13), respectively:

$$A_{ij} = \frac{b^{i+j+1} - a^{i+j+1}}{i+j+1}, i, j = 0,1,2\ldots,m \tag{12}$$

$$B_i = v_{Y_\alpha^U i}, i = 0,1,2\ldots,m \tag{13}$$

Based on the general case of the optimal square approximation method, the estimates of a and b are related to the skewness coefficients of Z_α^U:

$$a = \begin{cases} -3.5, v_{Y_\alpha^U 3} \geq 0 \\ -4.0, v_{Y_\alpha^U 3} < 0 \end{cases}, b = \begin{cases} 3.5, v_{Y_\alpha^U 3} \leq 0 \\ 4.0, v_{Y_\alpha^U 3} > 0 \end{cases} \tag{14}$$

Thereafter, $\lambda_i(i=0,1,\ldots,m)$ can be solved using Eq. (9), yielding the approximate polynomial expression $p_\alpha^U(z) = \sum_{i=0}^m \lambda_i z^i$. TOFM can be used when $m=4$.

4.3. TOFM based on envelope distribution using modern approximation method

As discussed above, based on the envelope distribution, which is a conservative description of fuzzy random variables, the fuzzy randomness problem is transformed into an approximate fitting problem on the interval of cut level. By employing the modern approximation method in TOFM, the approximate polynomial expression of the FPDF of limit state function can calculated.

It should be noted that due to the extension of the expansion in Eq. (3) to the second moment, only the first four moments of $v_{Y_\alpha^U i}, i = 0,1,2,3,4$ can be expressed. If higher central moments required, Pearson family curves could be used to develop the relationship between each central moment of the family curves, as follows:

$$v_{Y_\alpha^U i+1} = -\frac{k}{1+(k+2)c_2}\left[c_0 v_{Y_\alpha^U i-1} + c_1 v_{Y_\alpha^U i} \right], i = 4,5,\ldots \tag{15}$$

where $c_i, i = 1,2,3$ are the Pearson family curve parameters, which can be expressed in terms of the first four moments. It should be noted that the used constraint conditions order number is dependent on the specific case, and the order increases lead to an increase in the calculation time consumption. After obtaining the polynomial expression of $f_\alpha^U(z)$, the upper bound of the fuzzy reliability probability at the α cut level under different modern approximation method is:

$$(P_r)_\alpha^U \approx \Pr(Z_\alpha^U \geq 0) = \Pr(Y \geq -\frac{\mu_{Z_\alpha^U}}{\sigma_{Z_\alpha^U}}) = \begin{cases} \int_{-\infty}^{-\frac{\mu_{Z_\alpha^U}}{\sigma_{Z_\alpha^U}}} \sum_{i=0}^m \lambda_i y^i \, dy, \text{ optimal square approximation} \\ \int_{-\infty}^{-\frac{\mu_{Z_\alpha^U}}{\sigma_{Z_\alpha^U}}} \sum_{i=0}^m \exp\left(-\sum_{j=0}^m a_j y^j \right) dy, \text{ maximum entropy} \end{cases} \tag{16}$$

The lower bound of the reliability probability at the α cut level R_α^L is:

$$(P_r)_\alpha^L \approx \Pr(Z_\alpha^L \geq 0) = \Pr(Y \geq -\frac{\mu_{Z_\alpha^L}}{\sigma_{Z_\alpha^L}}) = \begin{cases} \int_{-\infty}^{-\frac{\mu_{Z_\alpha^L}}{\sigma_{Z_\alpha^L}}} \sum_{i=0}^m \lambda_i y^i \, dy, \text{ optimal square approximation} \\ \int_{-\infty}^{-\frac{\mu_{Z_\alpha^L}}{\sigma_{Z_\alpha^L}}} \sum_{i=0}^m \exp\left(-\sum_{j=0}^m a_j y^j \right) dy, \text{ maximum entropy} \end{cases} \tag{17}$$

Thus, the membership degree of reliability is obtained by performing the abovementioned process at each cut level. The procedure involved in the TOFM based on the envelope distribution method can be summarized as follows:

Step 1. The family distribution of the fuzzy random variables under the given cut level can be obtained according to the membership interval of the fuzzy random variables.

Step 2. The envelope distribution is constructed for each fuzzy random variable at each cut level using the method presented in Section 3.

Step 3. Based on the envelope distribution, the respective bound central moments $\nu_{x_\alpha^U i}, i = 0,1,2,3,4$ and $\nu_{x_\alpha^L i}, i = 0,1,2,3,4$ of $f_\alpha^U(x)$ and $f_\alpha^L(x)$ are obtained using a statistical method.

Step 4. The bounds of the limit state function $Z = g(\tilde{\mathbf{X}})$ at the α cut level Z_α^U and Z_α^L are approximated by the first two orders of the Taylor series expansion Each are normalized to Y_α^U and Y_α^L.

Step 5. The bound central moments $v_{Y_\alpha^U i}, i=0,1,2,3,4$ and $v_{Y_\alpha^L i}, i=0,1,2,3,4$ are calculated using Eq. (3), and the higher order moments are calculated using Eq. (15), if required.

Step 6. Employing the modern approximation method by considering the bound central moments as the constraint conditions. In particular, for the maximum entropy model, as mentioned in section 4.1, the approximate polynomial expression is obtained using Eq. (7). For the optimal square approximation method, as mentioned in Section 4.2, **A** and **B** are calculated using Eqs. (12) and (13), respectively. The fitting polynomial coefficients $\lambda_j, j=0,1,2,\ldots,m$ are then obtained using Eq. (9).

Step 7. The bounds of the reliability at the α cut levels $(P_r)_\alpha^U$ and $(P_r)_\alpha^L$ are calculated using Eqs. (16) and (17), respectively. The steps above can be repeated at each α cut level, and yielding the fuzzy reliability. Fig. 3 shows a flowchart of the proposed method.

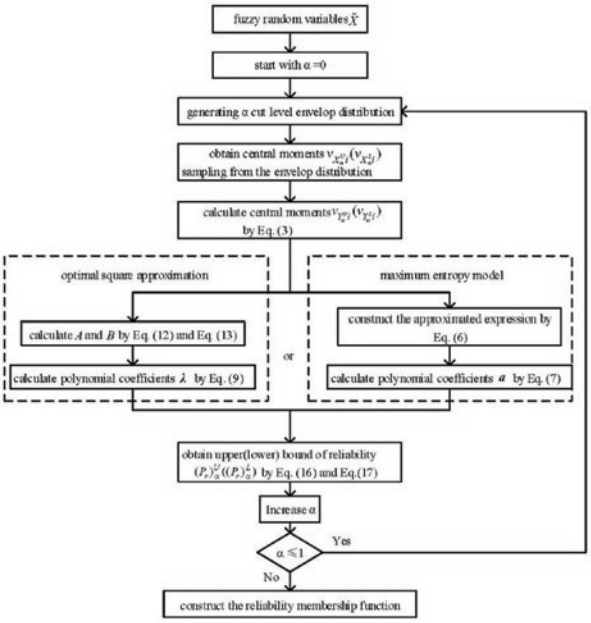


Fig. 3. Flowchart of TOFM based on the envelope distribution using the modern approximation method.

5. Examples

Three examples are presented to illustrate the proposed method. The first is a pure mathematical example. The second and third examples demonstrate the applicability of the proposed method in engineering, i.e., loads or materials that are considered with respect to fuzzy randomness uncertainty. Results from the MCS and FFORM methods are compared with those from the proposed method as these are classical approaches to fuzzy random uncertainties. The numerical results illustrate the superiority of the present approach in terms of efficiency and accuracy. The results contain sharp enclosures for all values of the reliability probability based on the proposed method compared with those obtained by MCS and FFORM approaches.

5.1. Investigation 1 (numerical)

It is assumed that the limit state function of the structure is $Z=x_1 * x_2 - x_3 - 1200$, where $x_1 \sim N\left(\tilde{\mu}_{x_1}, \tilde{\sigma}_{x_1}\right)$, $x_2 \sim N\left(\tilde{\mu}_{x_2}, \tilde{\sigma}_{x_2}\right)$, and $x_3 \sim N\left(\tilde{\mu}_{x_3}, \tilde{\sigma}_{x_3}\right)$. The basic realization of $x_i, i=1,2,3$ are assumed as normal distribution. The mean μ and standard deviation σ of the basic variables are considered as triangular fuzzy numbers:

$$\mu_{x_1} = 37.5;38;40 \ ; \ \sigma_{x_1} = 1.6;2;2.4$$

$$\mu_{x_2} = 53.5;54;56 \ ; \ \sigma_{x_2} = 3.6;4;4.4$$

$$\mu_{x_3} = 19.7;20;21 \ ; \ \sigma_{x_3} = 1;1.5;2$$

α discretization is used for mapping the fuzzy space to the interval random space. Moreover, α is varied from 0 to 1, and the fuzzy numbers are evaluated at the following α: 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0. At each level, an interval is obtained for each distribution parameter, and the entire support domain of the problem is obtained for α = 0. Permissible domains for the distribution parameters could be easily calculated for different values of α. The results from the MCS and FFORM methods are compared in this example, as shown in Fig. 4.

For the MCS method, 64 combinations are used at each α cut level ($N(\bar{\mu},\bar{\sigma}), \bar{\mu}=\left[\mu_{x_i}^{low}, \mu_{x_i}^{up}\right], \bar{\sigma}=\left[\sigma_{x_1}^{low}, \sigma_{x_1}^{up}\right], i=1,2,3$), and 1,000,000 analyses are required for each combination, thus the MCS method required 64 × 1,000,000 × 6(384,000,000) runs. For the FFORM method, four iterations are performed for constructing the bound fuzzy reliability index, which indicates that the FFORM method required 64 × 4 × 6 (1536) runs. In comparison, the proposed method required two repetitions of the process at the upper and lower boundaries. Moreover, the maximum error at a given α cut level is found to be 1.54×10^{-2} at α = 0 of the lower distribution, as shown in Table 1. This error level is acceptable compared to the entire reliability membership function. The proposed method provides a clear improvement in the calculation efficiency, and the results obtained by the three methods are similar. In addition, 1000 samples are used to construct the envelope distribution, thus the proposed method required 6 × 1,000 × 6 (36,000) sampling operations. A comparison of the computation time is shown in Table 1, which illustrate the great advantage of the conventional methods. The result of the proposed method is included in the MCS and FFORM methods as shown in Fig. 4, that's because the boundary extremum occurs when $N\left(\mu_{x_i}^{low}, \mu_{x_i}^{up}\right)$ and $N\left(\mu_{x_i}^{up}, \mu_{x_i}^{low}\right)$ are operated. This indicates that the proposed method has the effect of correcting and amplifying reliability when the extremum is conservative. This example demonstrates the superiority of TOFM based on the envelope distribution approach with respect to other approaches in the reliability assessment of structures. In the following two examples, the efficiency of the proposed method is illustrated based on evaluation.

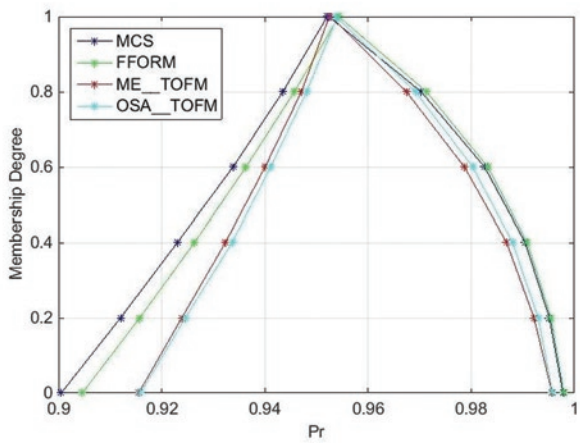


Fig. 4. Reliability membership function in Example 1

Table 1. Fuzzy reliability probability for the example 1

	ME_TOFM Reliability probability		OSA_TOFM Reliability probability		FFORM Reliability probability		MCS Reliability probability	
α	R_L	R_U	R_L					
1.0	0.95247	0.95279	0.95405	0.95372	0.95433	0.95433	0.95208	0.95241
0.8	0.94707	0.96744	0.94807	0.96920	0.94570	0.97134	0.94347	0.97030
0.6	0.94004	0.97868	0.94111	0.98028	0.93634	0.98320	0.93382	0.98258
0.4	0.93233	0.98687	0.93363	0.98802	0.92633	0.99088	0.92301	0.99049
0.2	0.92389	0.99218	0.92455	0.99297	0.91573	0.99547	0.91201	0.99517
0.0	0.91553	0.99569	0.91591	0.99578	0.90460	0.99796	0.90051	0.99780
Computation time	123.98 s		115.55 s		151.41 s		188.73 s	

5.2. Investigation 2 (model)

A roof truss is presented in Fig. 5, for which the top boom and compression bars are reinforced with concrete, and the bottom boom and the tension bars are made of steel. This evaluation is conducted under the assumption that a uniformly distributed load q is applied to the roof truss, and that a uniformly distributed load can be transformed into the nodal load $P = ql/4$. The perpendicular defection Δ_c of node C can be obtained through mechanical analysis, and it is a function of the basic variables, i.e., $\Delta_c = \frac{ql^2}{2}\left(\frac{3.81}{A_c E_c} + \frac{1.13}{A_s E_s}\right)$, where A_c, A_s, E_c, E_s, and l are the cross-sectional area, elastic modulus, length of concrete, and length of the steel bars. With respect to safety and applicability, the limit condition is that Δ_c of node C could not exceed 3.1 cm, and the limit state function could be constructed using $g = 0.031 - \Delta_c$. The values of l and q are shown in Table 2.

Table 2. Variables in Example 2

Variable	Value
$l(m)$	12
$q(10^4\ N)$	2

In this example, the basic realization of the fuzzy random variables A_C, A_S E_C, and E_S are assumed as the normal distribution. The mean and standard deviation of the variables are considered as triangular fuzzy numbers:

$$\mu_{A_C} = 3.85; 4.0; 4.1\times10^{-2} m^2 \ ;\ \sigma_{A_C} = 0.29; 0.32; 0.35\times10^{-2} m^2$$

$$\mu_{A_S} = 0.99; 1.0; 1.01\times10^{-3} m^2 \ ;\ \sigma_{A_C} = 0.055; 0.06; 0.065\times10^{-3} m^2$$

$$\mu_{E_C} = 1.64; 1.67; 1.7\times10^{10} Pa \ ;\ \sigma_{E_C} = 0.13; 0.14; 0.18\times10^{10} Pa$$

$$\mu_{E_S} = 0.90; 0.91; 0.92\times10^{11} Pa \ ;\ \sigma_{E_C} = 0.037; 0.04; 0.045\times10^{11} Pa$$

α is varied from 0 to 1, and the fuzzy numbers are evaluated at the following α: 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0. The results from the MCS and FFORM methods are compared in this example.

The MCS method require 256 combinations at each cut level in this example with 1,000,000 runs for each combination, i.e., the MCS method required 256 × 1,000,000 × 6 (1,536,00,000) runs. For the FFORM method, six iterations are performed for each combination,

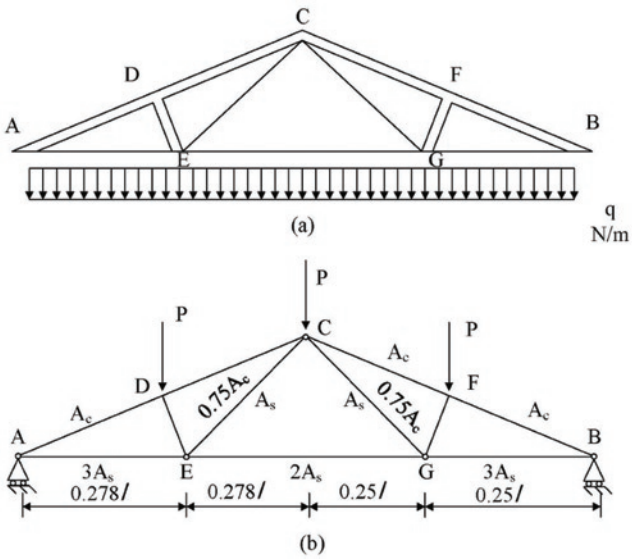


Fig. 5. Truss stress diagram

i.e., the FFORM method required 256 × 6 × 6 (9216) runs. In comparison, the proposed method requires two repetitions at the upper and lower boundaries. The maximum error for a given α cut level is 5.15×10^{-3} at Level 0, as shown in Table 3, which illustrates the reliability probability at each level using the three methods. In addition, 1000 samples are used to construct the envelope distribution, thus the proposed method required 8 × 1,000 × 6 (48,000) sampling operations, illustrating its advantage over the MCS method. In this case, the computational efficiency of the TOFM based on the envelope distribution is evident in this example compared with the results from the MCS and FFORM methods. As the number of dimensions and nonlinearity increase, the advantage of this method is demonstrated, as showed by computation time in Table 3. In contrast with Example 1, the result of the MCS and FFORM methods are included in the proposed method, that's because the proposed method boundary extremum occurs when $N\left(\mu_{x_i}^{low}, \mu_{x_i}^{low}\right)$ and $N\left(\mu_{x_i}^{up}, \mu_{x_i}^{up}\right)$ are operated. This indicates that the proposed method has the effect of correcting and diminishing reliability when the extremum occurs on the side combination instead of the cross combination of mean and standard deviation as shown in Example 1. The application of the proposed method in this paper to complex structures is presented below.

Table 3. Fuzzy reliability probability for Example 2

	ME_TOFM Reliability probability		OSA_TOFM Reliability probability		FFORM Reliability probability		MCS Reliability probability	
α	R_L	R_U	R_L	R_U	R_L	R_U	R_L	R_U
1.0	0.99780	0.99785	0.99796	0.99789	0.99769	0.99769	0.99765	0.99784
0.8	0.99617	0.99887	0.99591	0.99927	0.99637	0.99859	0.99643	0.99866
0.6	0.99403	0.99941	0.99350	0.99992	0.99447	0.99918	0.99466	0.99919
0.4	0.99078	0.99971	0.99007	0.99999	0.99186	0.99954	0.99224	0.99954
0.2	0.98595	0.99986	0.98522	0.99970	0.98837	0.99976	0.98867	0.99975
0.0	0.97929	0.99994	0.97934	0.99969	0.98383	0.99988	0.98444	0.99987
Computation time	136.2 s		139.31 s		187.04 s		239.45 s	

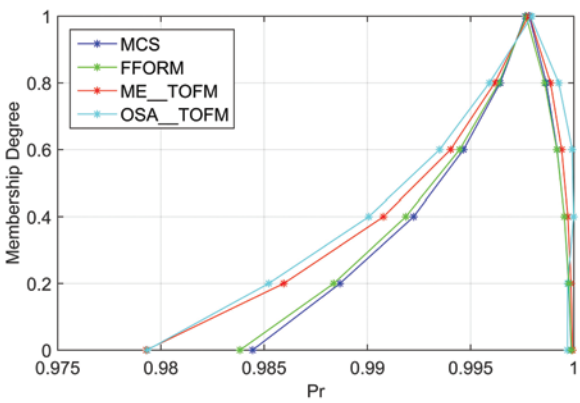


Fig. 6. Reliability membership function in Example 2.

5.3. Investigation 3 (model)

A truss member structural system is one of the most common structural forms in structural engineering. Fig. 7 shows the square grid structure of the square plate. The length of the upper chord plane is 5.0 m, the length of the lower chord plane is 4.0 m, the length of the string is 1.0 m, the height of the net frame is 0.7 m (the vertical distance between the upper and lower chords), the upper chord plane is hinged, and the lower chord are free. The bar is made of steel, and the mean values of the rod diameters are $4.91\times10^{-4}m^2$ with density of $7.8\times10^3 kg/m^3$. In addition, the mean modulus of elasticity and Poisson's ratio are 207 GPa and 0.3, respectively. Three loads labelled P44, P49, and P54(location node 44, 49 and 54) are considered as independent fuzzy random variables, as shown in Fig. 7. The loads are applied along the negative Z direction (Δ_Z). The serviceability limit state of the deflection is considered. The vertical deflection limit at any node is set as 4.57 cm. The limit state function could be constructed using $g = 4.57 - (\Delta_Z)_{max}$. The assumed diameter of the rods, modulus of elasticity, and Poisson's ratio are shown in Table 4.

The basic realization of the fuzzy random loads are assumed to be normal distribution. The mean and standard deviation are considered as triangular fuzzy numbers:

$$\mu_{P44} = 4800;5000;5200N \; ; \sigma_{P44} = 386;400;404N \; ;$$

Table 4. Variable value of Example 3

Variable	Value
Diameter of rods	4.91×10^{-4} m^2
Modulus of elasticity	207 GPa
Poisson ratio	0.3

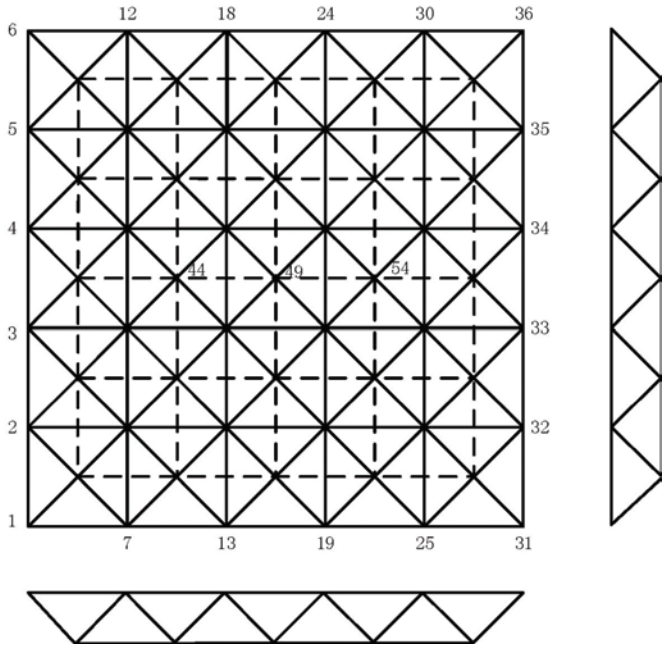


Fig. 7. Space-truss structure

$$\mu_{P49} = 9500;10000;10500N \; ; \sigma_{P49} = 792;800;808N \; ;$$

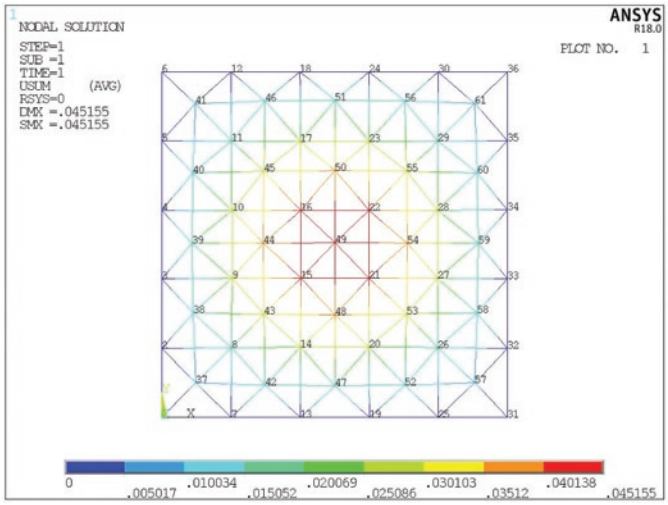
$$\mu_{P54} = 4800;5000;5200N \; ; \; \sigma_{P54} = 386;400;404N$$

α discretization is used to map this fuzzy space to the interval random space. Moreover, α is varied from 0 to 1 in intervals of 0.2. The results from MCS and FFORM approach are compared in this example.

The negative direction of the Z axis Δ_Z are calculated using finite element software ANSYS. Moreover, as a problem with implicit limit-state functions, multi-point approximations are constructed for the limit state, and the closed-form expressions could then be constructed to estimate the reliability bound. The Latin hypercube sampling technique is used to sample 35 design points in the abovementioned methods.

On this basis, the fuzzy reliability of the structure could be obtained using the proposed method, and the results for different cut levels are listed in Table 5. The maximum displacement along the Z axis is shown in Fig. 8. The results show that node 49 is the point where maximum displacement occurs. The reliability membership function is presented in Fig. 9.

a)



b)

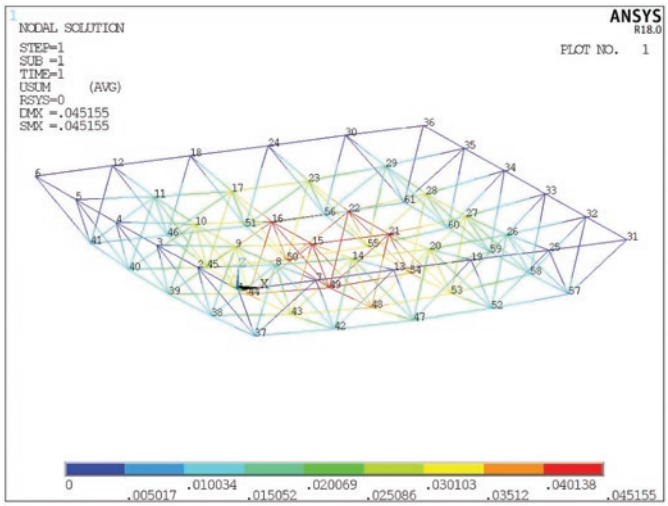


Fig. 8. Displacement diagram in ANSYS

Similarly, the MCS method had 64 combinations in this example, i.e., it required 64 × 1,000,000 × 6 (384,000,000) runs. For the FFORM method, seven iterations are used in the bound distribution, which required 64 × 7 × 6 (2688) runs. In comparison, the proposed method also requires two repetitions of the process at the upper and lower boundaries. Table 5 shows the reliability probability at each cut level. Moreover, 1000 samples are used to construct the envelope distribution. Therefore, the proposed method required 6 ×1,000 × 6 (36,000) samples. Table 5 shows a comparison of the computation time with each method. Fig. 9 shows that the membership function of the proposed method exhibits conservative characteristics compared with the MCS method, which is more precise than the FFORM method due to the increased nonlinearity. There is a clear increase in efficiency that is significant considering the structural complexity.

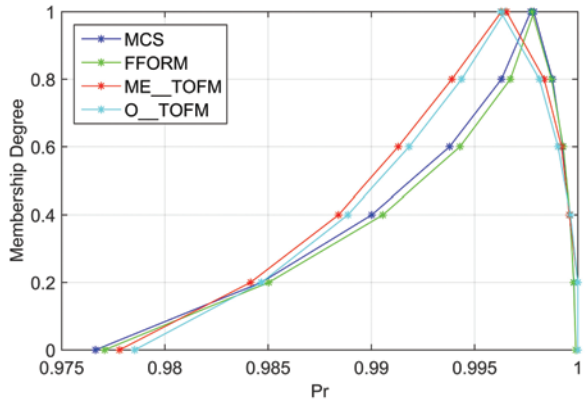


Fig. 9. Reliability membership function in Example 3

6. Conclusion

In this study, a novel structural reliability analysis method with an uncertainty information model is applied to fuzzy random variables. The fuzzy reliability is calculated by using TOFM based on the envelope distribution. In the proposed method, based on the conservative characteristics of the bound distribution, the envelope distribution is used to describe the fuzzy random variables, which converts the fuzzy randomness into a probability problem. Hence, the bounds of the fuzzy reliability are calculated. Without the requirement of an iterative algorithm for calculating the reliability index β, the proposed method provides a significant advantage with respect to the simplification of the reliability calculation and the increased efficiency of the reliability analysis.

As illustrated in the examples, by combining with the modern approximation method, the proposed method only requires the central moments of each variable, which eliminates numerous iterative processes. Moreover, the calculation scale is considerably reduced compared with conventional reliability analysis methods, which significantly broadens its applicability. As the number of uncertainty variables increases, the efficiency of the proposed method is significant when the performance of the compared methods is unsatisfactory. The results show that the proposed method has the correction function. The fuzzy reliability can be appropriately increased or decreases according to the combination of mean and standard deviation when extreme value occurs.

Table 5. Fuzzy reliability probability for the example 3

	ME_TOFM Reliability probability		OSA_TOFM Reliability probability		FFORM Reliability probability		MCS Reliability probability	
α	R_L	R_U	R_L	R_U	R_L	R_U	R_L	R_U
1.0	0.99630	0.99654	0.99647	0.99626	0.99788	0.99778	0.99772	0.99787
0.8	0.99390	0.99837	0.99437	0.99814	0.99673	0.99873	0.9963	0.99877
0.6	0.99130	0.99921	0.99183	0.99904	0.99429	0.99929	0.99379	0.99927
0.4	0.98841	0.99959	0.98886	0.99964	0.99056	0.99962	0.99002	0.99961
0.2	0.98415	1.0	0.98468	1.0	0.98503	0.9998	0.98467	0.99981
0.0	0.97780	1.0	0.97854	1.0	0.97708	0.9999	0.97664	0.99990
Computation time	115.92 s		115.22 s		238.74 s		276.43 s	

The results of proposed method tend to be conservative, and they are suitable for engineering applications. In particular, the central moments of the envelope distribution appropriately describe the upper and lower bounds of the numerical characteristics of fuzzy random variables, which can be used as the input in reliability or other analyses. Furthermore, several aspects can also be evaluated, i.e., how to quickly select the most suitable value of *m* and accurately estimating *a* and *b*.

Acknowledgement

his research was supported by the National Natural Science Foundation of China (No. 51675026).

References

1. Dempster A P. Upper and lower probabilities induced by a multivalued mapping. The Annals of Mathematical Statistics 1967; 38(2): 325-339, https://doi.org/10.1214/aoms/1177698950.
2. Dhande S G, Chakraborty J. Analysis and synthesis of mechanical error in linkages-a stochastic approach. Journal of Engineering for Industry 1973; 95: 672-676, https://doi.org/10.1115/1.3438208.
3. Du X. Reliability synthesis for mechanism. Machine Design 1996; 13(1): 8-11.
4. Gil M Á, Miguel L D, Ralescu D A. Overview on the development of fuzzy random variables. Fuzzy Sets and Systems 2006;157(19): 2546-2557, https://doi.org/10.1016/j.fss.2006.05.002.
5. Grandhi R V, Wang L. Higher-order failure probability calculation using nonlinear approximations. Computer Methods in Applied Mechanics & Engineering 2013; 168(1-4): 185-206, https://doi.org/10.1016/S0045-7825(98)00140-6.
6. Hryniewicz O. Bayes statistical decisions with random fuzzy data-an application in reliability. Reliability Engineering & System Safety 2016; 151: 20-33, https://doi.org/10.1016/j.ress.2015.08.011.
7. Huang H Z. Structural reliability analysis using fuzzy sets theory. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14: 284-294.
8. Jahani E, Muhanna R L, Shayanfar M A, Barkhordari M A. Reliability Assessment with Fuzzy Random Variables Using Interval Monte Carlo Simulation. Computer-Aided Civil and Infrastructure Engineering 2014; 29(3): 208-220, https://doi.org/10.1111/mice.12028.
9. Jiang C, Zhang Q F, Han X, Qian Y H. A non-probabilistic structural reliability analysis method based on a multidimensional parallelepiped convex model. Acta Mechanica 2014; 225(2): 383-395, https://doi.org/10.1007/s00707-013-0975-2.
10. Kato M, Ono T, Zhao Y G. Second-Order third-moment reliability method. Journal of Structural Engineering 2002;128(8):1087-1090, https://doi.org/10.1061/(ASCE)0733-9445(2002)128:8(1087).
11. Kiureghian A D, Ditlevsen O. Aleatory or epistemic? Does it matter? Structural Safety 2009; 31(2): 105-112, https://doi.org/10.1016/j.strusafe.2008.06.020.
12. Koç M L, Balas C E. Reliability analysis of a rubble mound breakwater using the theory of fuzzy random variables. Applied Ocean Research 2013; 39: 83-88, https://doi.org/10.1016/j.apor.2012.10.007.
13. Körner R. On the variance of fuzzy random variables. Fuzzy Sets and Systems 1997; 92(1): 83-93, https://doi.org/10.1016/S0165-0114(96)00169-8.
14. Kwakernaak H. Fuzzy random variables - I. Definitions and theorems. Information Sciences 1978; 15(1): 1-29, https://doi.org/10.1016/0020-0255(78)90019-1.
15. Laulin L, Kieffer M, Didrit O, Walter E. Applied interval analysis. Berlin: Springer, 2001, https://doi.org/10.1007/978-1-4471-0249-6.
16. Li D Q, Chen Y F, Lu W B. Stochastic response surface method for reliability analysis of rock slopes involving correlated non-normal variables. Computers and Geotechnics 2011; 38(1): 58-68, https://doi.org/10.1016/j.compgeo.2010.10.006.
17. Li D Q, Jiang S H, Cao Z J, Zhou W, Zhou C B, Zhang L M. A multiple response-surface method for slope reliability analysis considering spatial variability of soil properties. Engineering Geology 2015; 187: 60-72, https://doi.org/10.1016/j.enggeo.2014.12.003.
18. Li H Z, Chen F, Yang Z J, Wang L D, Kan Y N. Failure mode analysis on machining center based on possibility theory. 5th International Conference on Electrical Engineering and Automatic Control. Weihai. 2016: 627-636, https://doi.org/10.1007/978-3-662-48768-6_71.
19. Li H B, Nie X. Structural reliability analysis with fuzzy random variables using error principle. Engineering Applications of Artificial Intelligence 2018; 67: 91-99, https://doi.org/10.1016/j.engappai.2017.08.015.
20. Liu B. Uncertainty theory(4th ed.). Berlin: Springer, 2015, https://doi.org/10.1007/978-3-662-44354-5.
21. Liu Y B, Zhong Q, Wang G Y. Fuzzy random reliability of structures based on fuzzy random variables. Fuzzy Sets and Systems. 1997; 86(3): 345-355, https://doi.org/10.1016/S0165-0114(96)00002-4.
22. Lutterkort D, Peters J, Reif U. Polynomial degree reduction in the L2-norm equals best Euclidean approximation of Bézier coefficients. Computer Aided Geometric Design 1998;16(7): 607-612, https://doi.org/10.1016/S0167-8396(99)00025-4.
23. Möller B, Beer M. Engineering computation under uncertainty - Capabilities of non-traditional models. Computers & Structures 2008; 86(10): 1024-1041, https://doi.org/10.1016/j.compstruc.2007.05.041.
24. Möller B, Graf W, Beer M. Safety assessment of structures in view of fuzzy randomness. Computers & Structures 2003; 81(15): 1567-1582, https://doi.org/10.1016/S0045-7949(03)00147-0.
25. Möller B, Reuter U. Prediction of uncertain structural responses using fuzzy time series. Computers & Structures 2008; 86(10): 1123-1139, https://doi.org/10.1016/j.compstruc.2007.09.002.
26. Möller B, Reuter U. Uncertainty forecasting in engineering. Berlin: Springer, 2007.
27. Neumaier A. Interval methods for systems of equations. Cambridge: Cambridge University Press, 1990.
28. Penmetsa R C, Grandhi R V. Uncertainty propagation using possibility theory and function approximations. Mechanics Based Design of Structures & Machines 2003; 31(2): 257-279, https://doi.org/10.1081/SME-120020293.
29. Shapiro, Arnold F. Modeling future lifetime as a fuzzy random variable. Insurance: Mathematics and Economics 2013; 53(3): 864-870, https://doi.org/10.1016/j.insmatheco.2013.10.007.
30. Shi Z X, Yang X Q, Yang W, Cheng Q. Robust synthesis of path generating linkages. Mechanism and Machine Theory 2005;40(1):45-54.

https://doi.org/10.1016/j.mechmachtheory.2004.05.008
31. Smith S A, Krishnamurthy T, Mason B H. Optimized vertex method and hybrid reliability. 43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Denver. 2002: 1465, https://doi.org/10.2514/6.2002-1465.
32. Song J, Lu Z Z. Moment method for general failure probability with fuzzy failure state and fuzzy safety state. Engineering Mechanics 2008; 25(2): 71-77.
33. Terán P. Probabilistic foundations for measurement modelling with fuzzy random variables. Fuzzy Sets & Systems 2007; 158(9): 973-986, https://doi.org/10.1016/j.fss.2006.12.006.
34. Willner K, Möller B, Beer M. Fuzzy Randomness. Uncertainty in civil engineering and computational mechanics. Computational Mechanics 2005; 36(1): 83-83, https://doi.org/10.1007/s00466-004-0643-4.
35. Wang Z L, Huang H Z. An approach to system reliability analysis with fuzzy random variables. Mechanism and Machine Theory 2012; 52: 35-46, https://doi.org/10.1016/j.mechmachtheory.2012.01.007.
36. Xu W L, Zhang Q X. Probabilistic analysis and Monte Carlo Simulation of the kinematic error in a spatial linkage. Mechanism and Machine Theory 1989; 24(1): 19-27, https://doi.org/10.1016/0094-114X(89)90078-5.
37. Zavadskas E K, Vaidogas E R. Multiattribute selection from alternative designs of infrastructure components for accidental situations. Computer-aided Civil & Infrastructure Engineering 2010; 24(5): 346-358, https://doi.org/10.1111/j.1467-8667.2009.00593.x.
38. Zhang L, Sheng D, Dong Z X. Application of direct integration method in breakwater reliability analysis. Ocean Engineering 2011; 29(4): 103-107.
39. Zhao G T, Dong Y G, Song Z Y. Random reliability analysis based on the fuzzy theory. Journal of Hefei University of Technology 2010; 33(2): 249-253.
40. Zhao Y G, Ono T. A general procedure for first/second-order reliability method (FORM/SORM). Journal of Structural Engineering 1999; 21(2): 95-112, https://doi.org/10.1016/S0167-4730(99)00008-9.
41. Zheng J M, Wang G Z. Perturbing Bézier coefficients for best constrained degree reduction in the L 2 -norm. Graphical Models 2003; 65(6): 351-368, https://doi.org/10.1016/j.gmod.2003.07.001.

Lingfei YOU
Jianguo ZHANG
Qiao LI
Nan YE
School of Reliability and Systems Engineering
Beihang University
XueYuan Road No.37,HaiDian District, BeiJing, China

E-mails: youlinfei@buaa.edu.cn, zjg@buaa.edu.cn,
buaaliqiao@gmail.com, yenan@buaa.edu.cn

Article citation info:
MAIOR CBS, CHAGAS MOURA M, LINS ID. Particle swarm-optimized support vector machines and pre-processing techniques for remaining useful life estimation of bearings. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 610–619, http://dx.doi.org/10.17531/ein.2019.4.10.

Caio Bezerra Souto MAIOR
Márcio das CHAGAS MOURA
Isis Didier LINS

PARTICLE SWARM-OPTIMIZED SUPPORT VECTOR MACHINES AND PRE-PROCESSING TECHNIQUES FOR REMAINING USEFUL LIFE ESTIMATION OF BEARINGS

ZASTOSOWANIE MASZYN WEKTORÓW NOŚNYCH ZOPTYMALIZOWANYCH METODĄ ROJU CZĄSTEK ORAZ TECHNIK PRZETWARZANIA WSTĘPNEGO DO OCENY POZOSTAŁEGO OKRESU UŻYTKOWANIA ŁOŻYSK

The useful life time of equipment is an important variable related to system prognosis, and its accurate estimation leads to several competitive advantage in industry. In this paper, Remaining Useful Lifetime (RUL) prediction is estimated by Particle Swarm optimized Support Vector Machines (PSO+SVM) considering two possible pre-processing techniques to improve input quality: Empirical Mode Decomposition (EMD) and Wavelet Transforms (WT). Here, EMD and WT coupled with SVM are used to predict RUL of bearing from the IEEE PHM Challenge 2012 big dataset. Specifically, two cases were analyzed: considering the complete vibration dataset and considering truncated vibration dataset. Finally, predictions provided from models applying both pre-processing techniques are compared against results obtained from PSO+SVM without any pre-processing approach. As conclusion, EMD+SVM presented more accurate predictions and outperformed the other models.

Keywords: *big data, vibration signal, bearings, remaining useful life, empirical mode decomposition, wavelets transform, support vector machine, particle swarm optimization*

Okres użytkowania sprzętu jest ważną zmienną związaną z prognozowaniem pracy systemu, a możliwość jego dokładnej oceny daje zakładom przemysłowym znaczną przewagę konkurencyjną. W tym artykule pozostały czas pracy (Remaining Useful Life, RUL) szacowano za pomocą maszyn wektorów nośnych zoptymalizowanych rojem cząstek (SVM+PSO) z uwzględnieniem dwóch technik przetwarzania wstępnego pozwalających na poprawę jakości danych wejściowych: empirycznej dekompozycji sygnału (Empirical Mode Decomposition, EMD) oraz transformat falkowych (Wavelet Transforms, WT). W niniejszej pracy, EMD i falki w połączeniu z SVM wykorzystano do prognozowania RUL łożyska ze zbioru danych IEEE PHM Challenge 2012 Big Dataset. W szczególności, przeanalizowano dwa przypadki: uwzględniający kompletny zestaw danych o drganiach oraz drugi, biorący pod uwagę okrojoną wersję tego zbioru. Prognozy otrzymane na podstawie modeli, w których zastosowano obie techniki przetwarzania wstępnego porównano z wynikami uzyskanymi za pomocą PSO + SVM bez wstępnego przetwarzania danych. Wyniki pokazały, że model EMD + SVM generował dokładniejsze prognozy i tym samym przewyższał pozostałe badane modele.

Słowa kluczowe: *duże dane, sygnał drgań, łożyska, pozostały okres użytkowania, empiryczna dekompozycja sygnału, transformata falkowa, maszyna wektorów nośnych, optymalizacja rojem cząstek*

1. Introduction

System prognosis is a key factor within the condition-based maintenance (CBM) strategy and has been highlighted in different fields of science (Widodo and Yang [54]; Sutharssan *et al.* [47]). In this context, Remaining Useful Life (RUL) is a rather common measure used to characterize equipment performance (Sikorska, Hodkiewicz and Ma [46]). According to Si *et al.* [42], RUL is the useful life left at a particular time of operation, and is typically random and unknown. In fact, RUL is related with several factors (e.g. current degradation state, operating environment, system function) and should be estimated from available sources of information such as condition and health monitoring sensors. Even though there is no universally accepted best model to estimate RUL (Liao and Köttig [25]), current promising statistical methods have dealt with real-time big data (Bousdekis *et al.* [5]).

In fact, different signals can be collected in order to track the degradation of a system, and then build an accurate relationship between the current health condition state and RUL. Many signals (e.g. vibration, acoustic emission, temperature) can represent the evolution of degradation, and their analyses are as necessary as arduous (Chang *et al.* [7]; El-Thalji and Jantunen [13]; Ambhore *et al.* [2]). In this context, rotating equipment has received special attention due to its critical operating regimes, frequent failure modes and availability of measurements (e.g. vibration), allowing detection and isolation of incipient failures (Vachtsevanos *et al.* [54]).

Support Vector Machines (SVM) have been a successful technique for RUL estimation once it can deal with relative multi-dimensional datasets (Liu *et al.* [28]). Several SVM-based methods (Soualhi, Medjaher and Zerhouni [44]; Saha, Goebel and Christophersen [41]; Patil *et al.* [34]) have been proposed to predict RUL, taking into account that hybrid methodologies usually improve estimation accuracy

and overcome limitations of individual methods (Souto Maior *et al.* [48]).

However, SVM learning performance strongly depends on the quality of the input data. In fact, the direct use of the original series as input variables could consider irrelevant information (e.g. noise) and/or miss important features, which may generate imprecise predictions. Hence, specific techniques can be used as pre-processing tools in order to improve data input quality, and then obtain superior predictions from the learning method.

A notable pre-processing technique is Empirical Mode Decomposition (EMD), which decomposes the original series into a sum of Intrinsic Mode Functions (IMFs). According to Huang *et al.* [21], EMD is adaptive, empirical, direct and intuitive. Other specific pre-processing approach based on time-frequency analysis is Wavelet Transforms (WT). The idea behind WT is the same for the short-time fourier transform (Allen [1]), concentrating analysis on frequency filters. However, WT presents the best frequency/time resolution trade-off once it applies windows (filters) of various lengths.

Hence, this work proposes analyzing the ability of EMD-based models and WT-based models to correctly predict RUL when coupled with optimized-SVM. We compared and evaluated the prediction performance when applying both pre-processing techniques as well as predictions obtained without them. The big database considered was provided by FEMTO-ST Institute for the IEEE PHM 2012 Data Challenge focused on the estimation of the RUL for bearings (Nectoux *et al.* [35]) from vibration data.

The remainder of this article is organized as follows: Section 2 presents concepts and a theoretical background about rolling bearing and vibration signals, EMD, WT and SVM. Section 3 describes the methodology and steps adopted on the creation of models to estimate RUL of bearings. Section 4 presents the vibration big database and two cases in which the methodology was applied as well as the results of this application. Section 5 concludes remarks.

2. Theoretical background

2.1. Rolling Bearings and Vibration Signal

Rolling bearings are critical components of rotating machines and its fault diagnosis has subject of extensive research (Rai and Upadhyay [37]; Nikolaou and Antoniadis [33]). Generally, the main component considered on the analysis of localized defects in rolling bearings are the outer race, inner race, ball and cage (Prabhakar, Mohanty and Sekhar [38]).

Regarding to monitoring information (i.e. data), signals are broadly classified depending on the its specific type: vibration and acoustic, temperature and wear debris analysis (Tandon and Choudhury [52]). Particularly, vibration signals are a remarkable indicators for determining failure modes because they are easy-to-measure and provides adequate information, being commonly used in the condition monitoring and diagnosis of the rotating machinery (Chang *et al.* [7]; McKee *et al.* [30]). In this context, several standard vibration-based measures are commonly used for diagnosis purposes, including entropy, root mean square, signal amplitude, variance, kurtosis, as well as higher order statistics (Lybeck, Marble and Morton [29]).

In a fault state, vibration signals presents different pattern from healthy state, which allows failure identification (Chang *et al.* [7]). Indeed, localized faults in rolling bearing components produce a series of broadband impulse responses in the acceleration signals. Each component of the bearing rolling (e.g. outer and inner race; ball) has its own rotation frequency and wave behavior, which leads to a composed and complex signal (Randall and Antoni [41]) as depicted in Figure 1. Pre-processing techniques (e.g. EMD and WT) represents an alternative to deal with complex series creating a more manageable data, yet still carrying the important information.

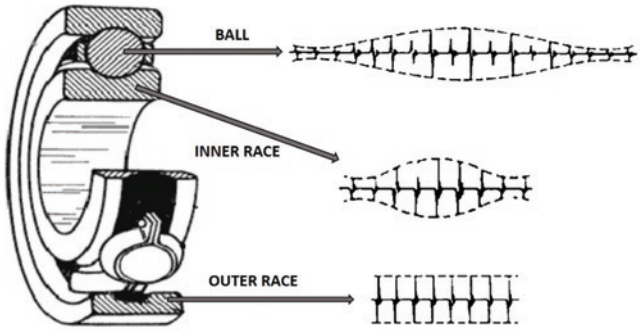


Fig. 1. Signals from local faults in rolling element bearings. Adapted from Randall and Antoni [38]

2.2. Empirical Mode Decomposition

A robust method to analyze non-linear and non-stationary data, Empirical Mode Decomposition (EMD) was developed by Huang *et al.* [19] and have been used in many types of applications. Its main idea is that a data series could be decomposed into a small number of simpler oscillation functions, called Intrinsic Mode Functions (IMFs). Then, the objective is to obtain IMFs regarding data characteristics in time scale (Huang and Wu [21]). Figure 2 depicts a general example of EMD decomposition (6 IMFs and a residue).

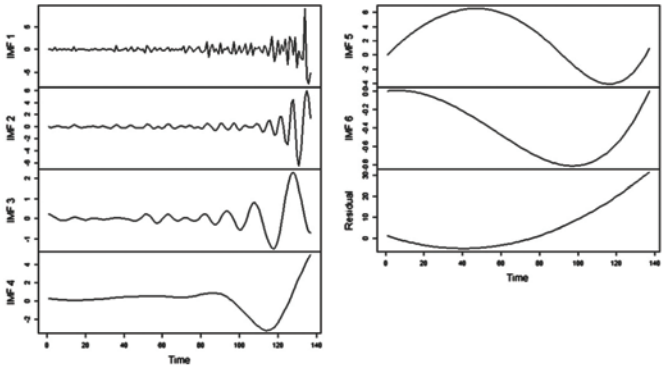


Fig. 2. General decomposition presented by EMD

Generally, any complex signal can be possibly separated into a small number of IMFs and a trend (or residue) r, indexed on $t \in T$, where T is the time interval (set of moments) considered. For a number N of IMFs, the original series $x(t)$ is expressed as follows:

$$x(t) = \sum_{i=1}^{N} \mathrm{IMF}_i(t) + r(t) \tag{1}$$

Huang *et al.* [19] defines IMF as a function that satisfies two conditions: (1) in the whole data set, the number of extrema and zero crossings must either equal or differ at most by one; and (2) at any point, the mean value of the envelop defined by the local maxima and the envelope defined by the local minima is zero. Then, EMD empirically identifies the IMFs through a process called sifting, which is based on three assumptions: (1) the signal has at least two extrema – one maximum and one minimum; (2) the characteristics time scale is defined by the time lapse between the extrema; and (3) if data has not extrema, but only contains inflection points, then it can be differentiated once or more times to reveal the extrema. The sifting goal is to remove riding waves to make the wave profile more symmetric. The sifting process can be described in the following steps (see. Figure 3):

1. Identify all local extrema (maximum and minimum) of the series $x(t)$;
2. Connect all the local extrema with a cubic spline line to create the upper and lower envelopes, $e_u(t), e_l(t)$, respectively;
3. Calculate the envelope mean $m(t) = \frac{e_u(t) + e_l(t)}{2}$
4. Obtain $h(t) = x(t) - m(t)$, which is candidate to be IMF;
5. Verify if $h(t)$ satisfies conditions defining an IMF. If it satisfies, an IMF was generated and the new series $x(t) - h(t)$ replaces the initial series $x(t)$. Otherwise, $h(t)$ would be processed again in step 1.

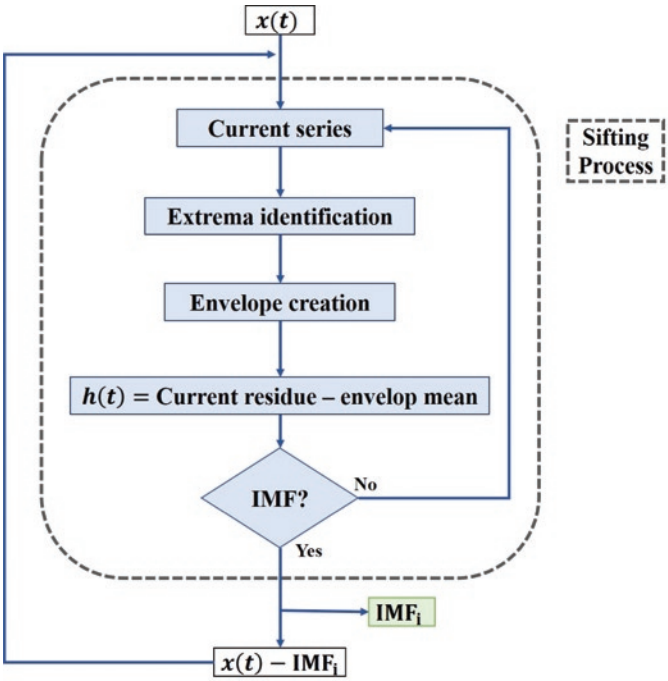


Fig. 3. Sifting process in EMD. Adapted from Souto Maior et al. [45]

At the end of the sifting process, a number of IMFs are generated as well as a final residue $r(t)$. The number of IMFs may vary depending on the intrinsic characteristics of $x(t)$. If the sifting process is carried to an extreme, the candidate IMF could have no physical meaning in sense of both amplitude and frequency modulations. Thus, a stop criterion for the sifting process has to be determined, which can be accomplished by limiting the standard deviation value computed from two consecutive sifting and/or the number of sifting iterations, as originally proposed by Huang *et al.* [19] and still in use (Eftekhar, Toumazou and Drakakis [13]). In practice, the number of IMFs created is lower than 10.

Generally, $\mathrm{IMF}_1(t)$ should contain the finest scale or the shortest period component of the signal. Since the reminder signal $r_1(t)$, i.e. $x(t) - \mathrm{IMF}_1(t)$, still contains information of longer periods (small frequencies), it is treated as the new data and it is subjected to the same sifting process as described above. This procedure can be repeated an all the subsequent iterations (Equation 2):

$$r_1(t) - IMF_2(t) = r_2(t), \cdots, r_{N-1}(t) - IMF_n(t) = r_N(t) \tag{2}$$

Finally, the original series $x(t)$ is represented as a sum of a number N of IMFs(t) and a residue $r_N(t)$, as presented in Equation (1).

2.3. Wavelet transform

Wavelet Transforms (WT) was first proposed by Morlet *et al.* [31] and has been a widespread technique applied in the field of signal analysis. WT is a mathematical tool that converts a signal of time domain using a wavelet basis function (i.e. a series of wavelet coefficients in time-scale domain) into a different form (Mallat [28]; Yan, Gao and Chen [57]). Kumar and Foufoula-Georgiou [22] remarks that a WT is chosen so that it has two important properties: admissibility (i.e. zero mean) and regularity (i.e. sufficient fast decay, to obtain localization) conditions.

The representation of the transform process occurs by an infinite series expansion of dilated/contracted and translated versions of a mother wavelet, each multiplied by an appropriate coefficient. Hence, the same signal could be represented in different forms, allowing multiple analysis. In practical applications, it is possible to use different well-known WT for distinct purposes and its choice depends on the specific signal characteristics. Figure 4 depicts a general signal processed by WT.

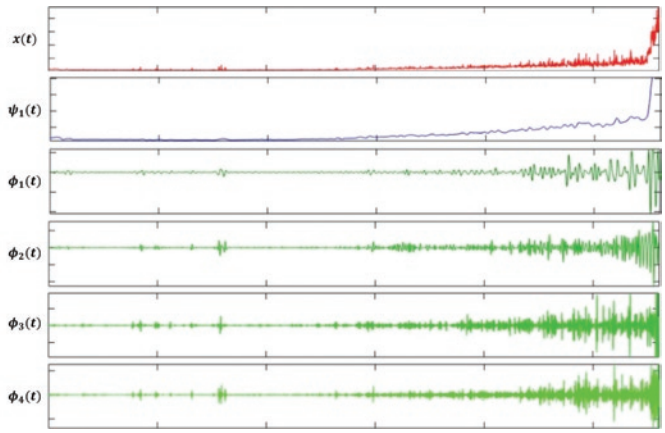


Fig. 4. General decomposition presented by WT

In time domain, a general wavelet dictionary $\{\psi_{u,s}\}$ can be defined as the dilated with the parameter $s > 0$, and translated by $u \in R$ of the mother wavelet ψ as follows(Chen *et al.* [8]):

$$\psi_{s,u}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \tag{3}$$

Hence, the WT of a function $x(t)$ is calculated by:

$$W(u,s) = \int_{-\infty}^{\infty} x(t) \psi_{s,u}(t) dt \tag{4}$$

Guohua *et al.* [17] argues that wavelet analysis decomposes a signal into two parts, called approximations and details, in which the former consists of high scale low frequency components and offers general information, while the latter corresponds to the low scale high frequency portions and provides detailed hidden patterns.

Daubechies [10], along with Mallat [30], popularized WT, allowing more liberty in the choice of the basis wavelet functions at a little expense of some redundancy, and is credited with the development of the wavelet from continuous to discrete signal analysis. Considering Equation (3), if s represents a continuous variable, then $W(u,s)$ is the continuous WT of $x(t)$ while if $s = a^i$, a is the scale parameter, then $W(u,s)$ is the discrete WT of $x(t)$ (Chen *et al.* [8]). Daubechies

wavelets basis relies on the scaling function $\phi(t)$, with set of (filter) coefficients $\{a_k\}_{k\in Z}$, and wavelets function $\psi(t)$, with set of (filter) coefficients $\{b_k\}_{k\in Z}$, satisfying the following refinement (Bakhoday-Paskyabi, Valinejad and Azodi [3]):

$$\phi_k(t) = \sqrt{2}\sum_k a_k \phi(2t-k) \tag{5}$$

$$\psi_k(t) = \sqrt{2}\sum_k b_k \phi(2t-k) \tag{6}$$

Hence, in WT decomposition, the discrete series $x(t)$ of M points is decomposed in distinct levels (e.g. j layer) of $\phi_k(t)$ and $\psi_{j,k}(t)$, each one related with a specific time-frequency characteristic, was follows (Chun-Lin [10]):

$$x(t) = \frac{1}{\sqrt{M}}\sum_k W(j_0,k)\phi_k(t) + \frac{1}{\sqrt{M}}\sum_j\sum_k W(j,k)\psi_{j,k}(t) \tag{7}$$

For discrete WT, Daubechies wavelets were used in this work due to its successful and acknowledged applications (Rafiee, Rafiee and Tse [36]; Genovese *et al.* [16]).

2.4. Support vector machine and particle swarm optimization

Support Vector Machine (SVM) is a supervised learning method which aims at create an mapping function between an input vector $\boldsymbol{x}$ and an output scalar y based on the training data set $D = \{(\boldsymbol{x}_1, y_1),\ldots(\boldsymbol{x_m}, y_m)\}$ (Wang [56]). The objective is to find the function $f(\boldsymbol{x})$ with the smallest penalization with respect to the deviation from the real data and, at the same time, as flat as possible. Depending on the nature of output y (i.e. whether binary/categorical or real numbers), SVM assess different learning problems: (i) classification problem, when dealing with categorical classes (e.g. heath state of a machinery); and (ii) regression problem, when dealing with quantitative and real-valued parameters (e.g. RUL estimation) (Lins *et al.* [26]).

SVM is based on the principle of the Structural Risk Minimization and its concepts are built on the Statistical Learning Theory (Vapnik [55]). This means to solve a convex and quadratic optimization problem in which the Karush-Kuhn-Tucker (KKT) condition are necessary and sufficient conditions to guarantee a global optimum. The goal is not to look for the perfect alignment between the function $f(\boldsymbol{x})$ and D, but the best representation for the mapping (i.e. a trade-off between the data fitness and the generalization ability to predict new data). The regression hyperplane equation is represented by:

$$f(\boldsymbol{x}) = \boldsymbol{w^T x} + b \tag{8}$$

with $\boldsymbol{x}$ expressing the input data, and $\boldsymbol{w^T}$ and b the coefficients to be estimated minimizing the following regularized risk function:

$$\min_{\omega,b} C\frac{1}{m}\sum_{i=0}^{m}\psi_\varepsilon(y_i, f_i) + \frac{1}{2}\boldsymbol{w^T w} \tag{9}$$

in which:

$$\psi_\varepsilon(y_i, f_i) = \begin{cases} |y_i - f_i| - \varepsilon & if \quad |y_i - f_i| \geq \varepsilon \\ 0 & otherwise \end{cases} \tag{10}$$

where y_i is the i-th real output (i.e. the original data) while f_i is the i-th estimated value. Equation (10) is known as the Vapnik's ε-insensitive loss function, which implies a non-penalization when the points are inside a tube with radius ε. Hence, ε measures the performance in the training process related to the first term of Equation (9). The second term of the same equation is used as a smoothness function of $f(\boldsymbol{x})$ and is related to the machine's capacity of generalization represented by $\boldsymbol{w^T w}$. Yet, C is a trade-off for penalization between the empirical risk and the model's smoothness.

In addition, the problem could be formulated using the primal-dual relation, which states that the solution from dual problem is also solution for the primal one. In practice, the dual problem is the one actually solved and, from the KKT conditions, a global solution is achieved. For more information related with the primal-dual problem, see Wright [55]. Hence, $f(\boldsymbol{x})$ is obtained in terms of the dual problem from Equation (8) as follows:

$$f(\boldsymbol{x}, \alpha, \alpha^*) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\boldsymbol{x}_i^T\boldsymbol{x} + b \tag{11}$$

where α_i and α_i^* are the dual Lagrange multipliers. To solve the linear regression, it is necessary to calculate the dot products, $\boldsymbol{x}_i^T\boldsymbol{x}$, $i = 1,2,\ldots,l$. The generalization for non-linear regression is possible by using kernel functions, $K(\boldsymbol{x}_i, \boldsymbol{x})$, $i = 1,2,\ldots,l$. Hence, Equation (11) becomes Equation (12):

$$f(\boldsymbol{x}, \alpha, \alpha^*) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(\boldsymbol{x}_i, \boldsymbol{x}) + b \tag{12}$$

We here adopted the gaussian Radial Basis Function (RBF) as the kernel function, which is expressed by $K(x_i, x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right)$, where γ is also a model parameter. One of several advantages of RBF over others kernel functions is to provide great flexibility requiring just one parameter (Lins *et al.* [26]).

A considerable challenge is to provide the best set of parameters to be used in training step. Therefore, metaheuristics, such as Particle Swarm Optimization (PSO), may lead to satisfactory parameters' values. PSO is a probabilistic optimization heuristic inspired by the social behavior of biological organisms (e.g., birds and fishes) and on the ability of animal groups to work as a whole in order to find some desirable position. This seeking behavior artificially modeled by PSO provides useful results in the quest for solutions of non-linear optimization problems in a real-valued search space (Bratton and Kennedy [6]). PSO-optimized SVM have been successfully applied in reliability problems (Droguett *et al.* [11]; Lins, Moura and Droguett [25]; García Nieto *et al.* [15]) and thus were here adopted for choosing to enhance models performance.

3. Methodology

The methodology proposed in this paper is presented in Figure 5 and was applied to an public bigdata set provided by FEMTO-ST Institute (Nectoux *et al.* [35]). The data was generated in the IEEE PHM 2012 Data Challenge focused on the estimation of the RUL for bearings based on vibration signals. Further details about the data set are exposed in next section.

The big datasets contain large quantity of information and, due to the computational cost and hardware restrictions, the learning model

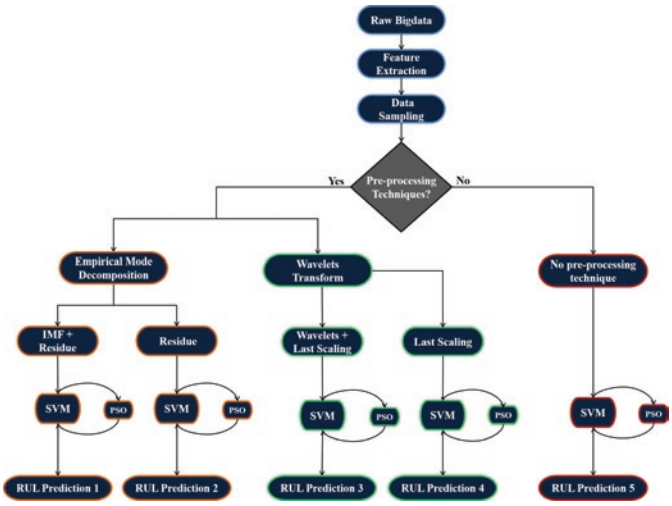


Fig. 5. Methodology applied for RUL prediction

cannot directly handle such an extensive data. Hence, to cover the massive data, two previous steps were performed for data dimension reduction: (i) feature extraction and (ii) data sampling. The former aims to reduce 2,560 vibration signal points into a representative measure (e.g. mean, kurtosis, or the highest absolute value), while the latter consists in sampling from original data (e.g. with frequency rate depending on the degradation state). Indeed, for healthier states of bearing, lower sampling frequency is necessary, while for more degraded states, higher sampling frequency is required. The procedures described above intended to extract only substantial information to be handled by the pre-processing techniques (EMD/WT).

After sampling, EMD or WT was performed. In each case, two distinct regression models were created. For EMD, one model contained all IMFs and the residue, while the other model contained just the final residue. For WT, one model consisted in wavelet functions of each level and the last scaling function, while the other model consisted in just the last scaling function. Given that SVM highly depends on the data input, the idea of using just the final residue (EMD-case) and the last scaling function (WT-case) was to provide an input possibly smooth enough yet still carries valuable aspects of the signal. For EMD, the number of IMFs generated directly depends on the characteristics of signal; a maximum tolerance of 20 sifting was used. In WT, Daubechies function was used as mother wavelet and 4 decomposition levels were applied.

The next step was to input PSO+SVM model with the previous processed data. Then, we evaluated each model (1 – IMFs + Residue; 2 – Residue; 3 – Wavelets + Scaling; 4 – Scaling; 5 – No pre-processing) based on the performance of RUL prediction. This methodology was applied in two different cases. The first application was performed with the complete data set provided. In this case, we considered data until failure (red dotted line in Figure 6), and a regression model was created to estimate the RUL considering each data point. The second application was more challenging once only part of the test set is provided, i.e., there was vibration signals just until some point far from failure time (yellow dotted line in Figure 6). In this case, the goal is to estimate correctly the RUL based on the current behavior of the vibration signals. Further details about both cases are presented in next session.

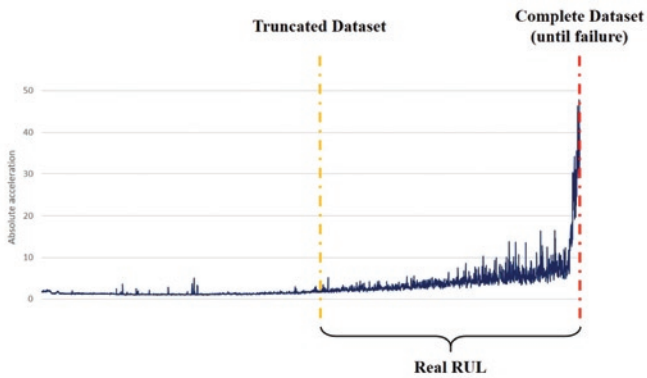


Fig. 6. Different cases considered on the test phase of application example

4. Application example

The presented methodology was applied to a real bigdata set provided by FEMTO-ST Institute (Nectoux *et al.* [35]). Experiments were carried out on a laboratory experimental platform (PRONOSTIA), that enables accelerated degradation of bearings under constant and/or variable operating conditions, while gathering online health monitoring data (e.g. vibration). The main objective is to provide experimental data that characterize the degradation of ball bearings along their complete operational life (until their total failure). Yet, considering the nature of a PHM challenge, data was complex and tricky, which really jeopardize the prediction capacity of the proposed models. The database have become popular during recent years, however many applications only use the complete dataset (Ren *et al.* [39]; Fumeo, Oneto and Anguita [14]) or do not reproduce the design and metrics of the Challenge (Boškoski *et al.* [4]; Mao *et al.* [29]) which is done in this work. For further information, see Nectoux *et al.* [32].

In our applications, we divided the dataset into training and test groups, where the former is necessary to teach SVM about the bearing degradation behavior, while the latter tries to predict correctly the behavior of an unseen bearing. The IEEE PHM Data Challenge provided one set of vibration data from a bearing to be used in the training phase, which is here called as 'Training Bearing', and had 2,803 observations in a run-to-failure experiment. Based on its behavior, estimations for RUL should be performed for another bearing (i.e. the test phase), here named as 'Test Bearing'. Finally, comparison between predictions obtained from models with EMD, WT and without a pre-processing technique are made in order to identify the most suitable approach.

SVM supervised learning method requires both y (i.e. the response variable) and $\boldsymbol{x}$ (i.e. the regressor/input) variables. In all cases, the response variable was the RUL and the regression variables were the vibrations signal. As previously mentioned, in EMD case, two models were created: one considering each IMF and the residue as regressors, and the other considering only the residue as the regressor. In WT case, two other models were also created: one containing each Wavelet and the last Scaling function as regressors, and the other considering only the last scaling function. The last model, without EMD or WT as pre-processing techniques, had the direct signal used as regressor.

In both application cases, it is not expected the direct point prediction to be enough precise due to the high variability of the data. However, the trend of all predictions should express the realistic RUL

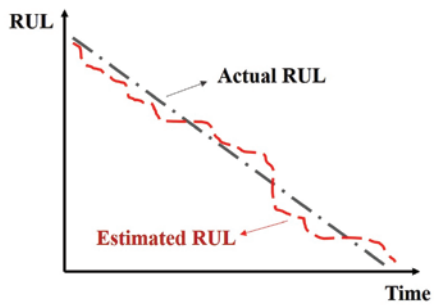


Fig. 7. Expected estimated RUL behavior. Adapted from Sutrisno et al. [48]

estimation, as seen in Figure 7 (Sutrisno *et al*. [51]). Thus, we expect the final residue for EMD and/or the last scaling function for WT to provide interesting results since both have intrinsic attributes related with flatness and physical meaning of the signal trend.

The bigdata set provided for 'Training Bearing' had 2,803 recordings, each one containing 2,560 points for horizontal vibration and other 2,560 points for vertical vibration. Hence, more than 14 million points were provided only for training purposes. In our analysis, a third signal composed by the vectorial sum of the horizontal vibration and vertical vibration was also computed. For each of the three vibration signals (i.e. vertical, horizontal and vectorial sum), three metrics were calculated in the feature extraction step: absolute peak amplitude, kurtosis and entropy. Figure 8 summarizes the data provided and the feature extraction process for the large amount of information.

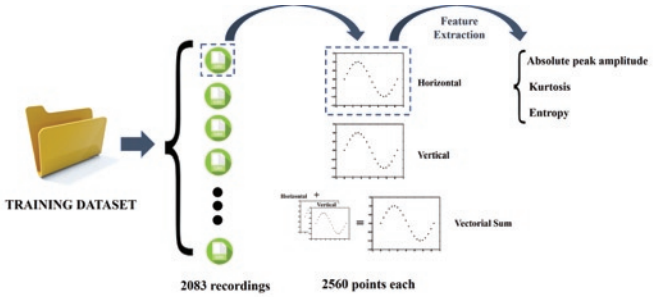


Fig. 8. Feature extraction process considering the dataset

By conducting several investigations and tests, the absolute peak amplitude proved to be the most suitable feature to be used in the next steps, similarly as in other papers (e.g. Rohlmann *et al.* [40]; Chen *et al.* [9]). Specifically, the horizontal vibration signal presented better initial results and was the chosen one to be analyzed. Moreover, dealing with the absolute amplitude, we considered the average of the five highest absolute peak acceleration values measured in each observation (archive). Averaging was done in order to alleviate the effect of data noise(Lee and Yun [24]).

After feature extraction, the data were categorized in four different regions in similar procedure to ISO 10816 (Standardization [49]) that deals with condition monitoring based on vibration. Therefore, each region represents a degradation phase of the bearing. We considered that a change of region takes place when the vibration trend line for the current region suffers a sudden increase of inclination (e.g. a new crack appears). In order to reduce the amount of information, a

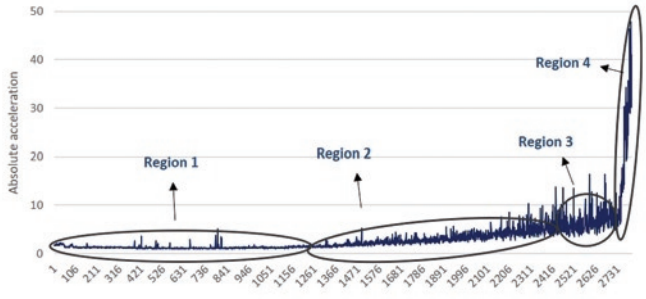


Fig. 9. The four different regions of degradation

Table 1. Sampling frequency and duration for each degradation region

Degradation Region	Sampling Frequency (in seconds)	Total time (in seconds)	Number of points considered
1	400	12000	30
2	200	13000	65
3	100	2500	25
4	0	510	17

data sampling was performed in every region with distinct sampling frequency (i.e. the more unstable the bearing is, the more necessary is to monitoring). Table 1 depicts the sampling frequency, the total duration for each degradation region and the number of points actually used after sampling. To illustrate, the third region was sampled every 100 seconds during the 2500 seconds in which the bearing stayed in this region, providing a total of 25 points. The four regions are shown in Figure 9.

4.1. Complete Dataset (until Failure)

In the first case, estimations of RUL for 'Test Bearing' were performed for all data, i.e. every test point until failure has an estimated RUL. As previously mentioned, it is not expected a good punctual prediction, but the trend should correctly express the degradation behavior. Hence, each model provides its own point prediction and a linear trend is created based on these points (i.e. take every predicted value from specific model and calculate the linear regression model representing them). This procedure was performed for the five models under analysis: 1) IMFs + residue; 2) residue; 3) Wavelets + last scaling; 4) Last scaling; and 5) no pre-processing.

In order to measure the quality of the estimated RUL, the Absolute Percentage Error (APE) was calculated, quantifying the distance error from the real RUL to the estimated one. Table 2 presents the APE as well as the number of support vectors related to each model. As it can be seen, EMD-based models (Model 1 and Model 2) presented superior performance compared with WT-based models and the model with no pre-processing technique. Moreover, even the worst EMD-based model was almost three times better than the others. As one might expected, Model 2, which use only the Residue as regressor, clearly presented the best performance.

Table 2. Errors for all tested models

Model	Regressors	Number of support vectors	APE
1	IMFs + Residue	62	2.54%
2	Residue	120	1.45%
3	Wavelets + Last Scaling	126	8.70%
4	Last Scaling	117	15.00%
5	Direct Vibration Data	100	7.58%

4.2. Truncated Dataset (IEEE PHM Data Challenge)

The second case aims to replicate exactly the IEEE PHM 2012 Data Challenge, which presents unclear end-of-life signature and unbalanced dataset (Huang *et al.* [18]). In this case, only truncated data was provided for 'Test Bearing' and the challenge was to estimate the actual RUL based on its initial degradation behavior (see Figure 6). All procedures applied in the first case was done: feature extraction and sampling to reduce the amount of data, training with 'Training Bearing' and test with 'Test Bearing' truncated data (i.e. not data until failure).

Vibration data provided for 'Test Bearing' consisted of 1,802 records and is depict on Figure 10. It is expected the bearing to pass through all four degradations regions, even if the truncated data does not present all of them. By analyzing, 'Test Bearing' does not present several abrupt changes in the signal, seemingly representing only first and second degradation regions. Therefore, inference had to be done to further degradation zones.

Analogously to session 4.1, the first region of 'Test Bearing' lasted 12,000 seconds. Figure 10 shows that vibration from the healthier stage (i.e. first degradation

region) is almost stationary, with negligible fluctuations, even though it represents most of the data. In addition, for 'Test Bearing', all remaining points belonged to the second region. Note that there is one outlier near the end of truncated data (even if its value is far from the failure vibration value presented 'Training Bearing' – around 47 units), but this does not correspond to a trend change and it was probably due to a noise in test. Thus, this was still considered belonging to second region. Given that, using data from region 1 will not represent any gain about bearing degradation. Moreover, using data from region 1 will only deviate the overall trend.

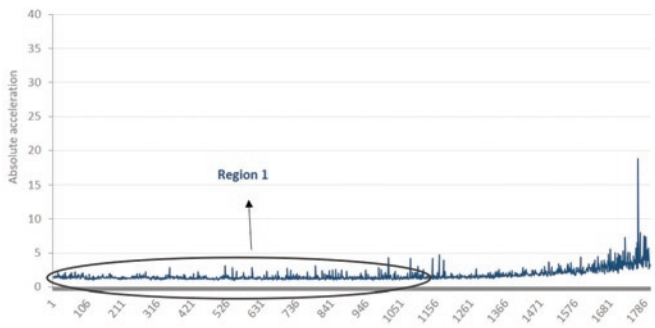


Fig. 10. Truncated test set used in second application

Therefore, in this case, none data from the first region was used. Also, sampling rate of the second region was adapted to provide more information, i.e. frequency of the fourth region was used (every 30 seconds). Thus, data test was reduced to nearly 200 test points used for estimation. Again, the same concept of first case was applied, with the predicted RUL being based on the overall trend of predictions. To define RUL estimation for the challenge, the trend line was extrapolated until it crosses the y -axis (i.e. trend estimate RUL equals to 0). Table 3 presents the models' performance based on the APE of the RUL estimation and real value of RUL. Errors for WT-models are not displayed, once they predicted negative values for RUL, which does not occur in reality.

Table 3. Errors for all tested models

Model	Regressors	APE
1	IMFs + Residue	15.39%
2	Residue	24.90%
5	Direct Vibration Data	58.53%

EMD-based models (Model 1 and 2), which also presented good performance in the first case, actually predicted best results. However, it is important to highlight the difficulty of this challenge, which is evident by considering the magnitude of errors in Table 3. Indeed, we compared our models with the winner of the challenge, which, based on the same evaluation metric, presented prediction errors of 37% (Sutrisno *et al.* [51]). The winner's prediction is worse than estimations provided by two of our models. Moreover, our best EMD-model (Model 1) reduced the error in more than 58%, which confirms the advantage in use the pre-processing method proposed.

5. Concluding remarks

This work compares the use of pre-processing techniques (i.e. EMD and WT) in order to increase prediction performance of RUL in PSO+SVM-based models. The comparison was applied to a real big data set of vibration signals of rolling bearings provided by an IEEE PHM Challenge competition. Two cases were performed: 1) considering the complete dataset (until failure); and 2) considering truncated (replicating the IEEE challenge). Even though performing alone PSO+SVM learning algorithm already provides reasonable estimations, applying pre-processing techniques yields gain in terms of prediction performance

Specifically, EMD based models (Model 1 and Model 2) presented the best performance compared with other approaches for both cases. Moreover, for the second case, two models provided better RUL predictions than the winner of the PHM Challenge competition. For future research, investigations about variations on the Wavelets approach (e.g. number of layers, type of mother wavelet) should be analyzed to improve its poor performance. Moreover, comparison with variants of EMD techniques (e.g. Ensemble Empirical Mode Decomposition (EEMD) (Wu and Huang [59]), Complete Ensemble Empirical Mode Decomposition (CEEMD) (Torres *et al.* [53]) could be done to verify if an even better prediction is achieved.

6. Acknowledgments

The authors thank the Brazilian research funding agencies 'Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)' and 'Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)' - Finance Code 001 - for the financial support through research grants.

References

1. Allen J. Short term spectral analysis, synthesis, and modification by discrete Fourier transform. IEEE Transactions on Acoustics, Speech and Signal Processing 1977; 25(3): 235-238, https://doi.org/10.1109/TASSP.1977.1162950.
2. Ambhore N, Kamble D, Chinchanikar S, Wayal. V. Tool condition monitoring system: A review. Materials Today: Proceedings 2015; 2(4-5): 3419-3428, https://doi.org/10.1016/j.matpr.2015.07.317.
3. Bakhoday-Paskyabi M, Valinejad A, Azodi H. D. Numerical solution of regularised long ocean waves using periodised scaling functions. Pramana 2019; 92(5): 71, https://doi.org/10.1007/s12043-019-1726-2.
4. Boškoski P, Gasperin M, Petelin D, Juricic D. Bearing fault prognostics using Rényi entropy based features and Gaussian process models, Mechanical Systems and Signal Processing 2015; 52-53: 327-337, https://doi.org/10.1016/j.ymssp.2014.07.011.
5. Bousdekis A, Magoutas B, Apostolou D. Mentzas G.Review, analysis and synthesis of prognostic-based decision support methods for condition based maintenance. Journal of Intelligent Manufacturing 2015; 29(6) 1303-1316, https://doi.org/10.1007/s10845-015-1179-5.
6. Bratton D. Kennedy J. Defining a Standard for Particle Swarm Optimization. 2007 IEEE Swarm Intelligence Symposium 2007; 120-127, https://doi.org/10.1109/SIS.2007.368035.
7. Chang L, Chung Y, Lin C, Chen J, Kuo C, Chen S. Mechanical Vibration Fault Detection for Turbine Generator Using Frequency Spectral Data and Machine Learning Model : Feasibility Study of Big Data Analysis. Sensors and Materials 2018; 30(4): 821-832, https://doi.org/10.18494/SAM.2018.1783.

8. Chen J, Li Z, Pan J, Chen G, Zi Y, Yuan J, Chen B, He Z. Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review. Mechanical Systems and Signal Processing 2016; 70-71: 1-35, https://doi.org/10.1016/j.ymssp.2015.08.023.
9. Chen X, Ding M, Wang T, Ding M, Wang J, Chen J, Yan J. Analysis and prediction on the cutting process of constrained damping boring bars based on PSO-BP neural network model. Journal of Vibroengineering 2017; 19(2): 878-893, https://doi.org/10.21595/jve.2017.18068.
10. Chun-Lin L. A Tutorial of the Wavelet Transform. Taipei: National Taiwan University, 2010.
11. Daubechies I. Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics 1993; 666-669, https://doi.org/10.1137/1.9781611970104.
12. Droguett E, Lins I, Moura M, Zio E, Jacinto C. Variable selection and uncertainty analysis of scale growth rate under pre-salt oil wells conditions using support vector regression. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2014; 229(4): 319-326, https://doi.org/10.1177/1748006X14533105.
13. Eftekhar A, Toumazou C, Drakakis E. M. Empirical Mode Decomposition: Real-Time Implementation and Applications. Journal of Signal Processing Systems 2013; 73(1): 43-58, https://doi.org/10.1007/s11265-012-0726-y.
14. El-Thalji I, Jantunen E. A summary of fault modelling and predictive health monitoring of rolling element bearings. Mechanical Systems and Signal Processing 2015; 60: 252-272, https://doi.org/10.1016/j.ymssp.2015.02.008.
15. Fumeo E, Oneto L, Anguita D. Condition based maintenance in railway transportation systems based on big data streaming analysis. Procedia Computer Science 2015; 53: 437-446, https://doi.org/10.1016/j.procs.2015.07.321.
16. García Nieto P. J, García-Gonzalo E, Sánchez Lasheras F, Juezc de Cos. Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability. Reliability Engineering and System Safety 2015; 138: 219-231, https://doi.org/10.1016/j.ress.2015.02.001.
17. Genovese L. Videau V, Ospici M, Deutsch T, Goedecker S, Méhaut J. Daubechies wavelets for high performance electronic structure calculations: The BigDFT project. Comptes Rendus Mécanique 2011; 339: 149-164, https://doi.org/10.1016/j.crme.2010.12.003.
18. Guohua G, Yu Z, Guanghuang D, Yongzhong Z. Intelligent Fault Identification Based On Wavelet Packet Energy Analysis and SVM. International Conference on Control, Automation, Robotics and Vision 2006; 1(3): 1-5, https://doi.org/10.1109/ICARCV.2006.345306.
19. Huang B, Jin C, Di Y, Lee J. Review of Data-Driven Prognostics and Health Management Techniques: Lessions Learned From Phm Data Challenge Competitions. Machine Failure Prevention Technology 2017.
20. Huang N. E, Shen Z, Long S, Wu M, Shih H, Zheng Q, Yen N, Tung C, Liu H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 1998; 903-995, https://doi.org/10.1098/rspa.1998.0193.
21. Huang N. E, Wu Z. A review on Hilbert-Huang transform: Method and its applications to geophysical studies. Reviews of Geophysics 2008; 46(2): 1-23, https://doi.org/10.1029/2007RG000228.
22. Huang S, Chang J, Huang Q, Chen Y. Monthly streamflow prediction using modified EMD-based support vector machine. Journal of Hydrology 2014; 511: 764-775, https://doi.org/10.1016/j.jhydrol.2014.01.062.
23. Kumar P, Foufoula-Georgiou E. Wavelet analysis for geophysical applications. Reviews of Geophysics 1997; 35(4), https://doi.org/10.1029/97RG00427.
24. Lee J. J, Yun C. B. Damage diagnosis of steel girder bridges using ambient vibration data. Engineering Structures 2006, https://doi.org/10.1016/j.engstruct.2005.10.017.
25. Liao L, Köttig F. Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. IEEE Transactions on Reliability 2014, https://doi.org/10.1109/TR.2014.2299152.
26. Lins I, Araujo M, Moura M, Silva M, Droguett E. Prediction of sea surface temperature in the tropical Atlantic by support vector machines. Computational Statistics and Data Analysis 2013; 61: 187-198, https://doi.org/10.1016/j.csda.2012.12.003.
27. Lins I, Moura M, Droguett E. Failure prediction of oil wells by support vector regression with variable selection, hyperparameter tuning and uncertainty analysis. Chemical Engineering Transactions 2013; 33: 817-822.
28. Liu Z, Wang L, Zhang Y, Chen C. A SVM controller for the stable walking of biped robots based on small sample sizes. Applied Soft Computing 2016; 38: 738-753, https://doi.org/10.1016/j.asoc.2015.10.029.
29. Lybeck N, Marble S, Morton B. Validating Prognostic Algorithms: A Case Study Using Comprehensive Bearing Fault Data, Aerospace Conference 2007; 1-9, https://doi.org/10.1109/AERO.2007.352842.
30. Mallat S. A Wavelet Tour of Signal Processing. A Wavelet Tour of Signal Processing 2009.
31. Mallat S. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 1989, https://doi.org/10.1109/34.192463.
32. Mao W, He J, Tang J, Li Y. et al. Predicting remaining useful life of rolling bearings based on deep feature representation and long short-term memory neural network. Advances in Mechanical Engineering 2018; 10(12), https://doi.org/10.1177/1687814018817184.
33. McKee K. K, Forbes G, Mazhar I, Entwistle R, Hodkiewicz M, Howard I. A vibration cavitation sensitivity parameter based on spectral and statistical methods. Expert Systems with Applications 2015; 42(1): 67-78, https://doi.org/10.1016/j.eswa.2014.07.029.
34. Morlet J, Arens G, Fourgeau E, Giardet D. Wave propagation and sampling theory-Part II: Sampling theory and complex waves. Geophysics 1982; 47(2): 222-236, https://doi.org/10.1190/1.1441329.
35. Nectoux P, Gouriveau R, Medjaher K, Ramasso E, Chebel-Morello B, Zerhouni N, Varnier C. PRONOSTIA : An experimental platform for bearings accelerated degradation tests. IEEE International Conference on Prognostics and Health Management 2012; 1-8.
36. Nikolaou N. G, Antoniadis I. A. Rolling element bearing fault diagnosis using wavelet packets NDT & E International 2002; 35(3): 197-205, https://doi.org/10.1016/S0963-8695(01)00044-5.
37. Patil M. A, Tagade P, Hariharan K, Kolake S, Song T, Yeo T, Doob S. A novel multistage Support Vector Machine based approach for Li ion battery remaining useful life estimation. Applied Energy 2015; 159: 285-297, https://doi.org/10.1016/j.apenergy.2015.08.119.
38. Prabhakar S, Mohanty A. R, Sekhar A. S. Application of discrete wavelet transform for detection of ball bearing race faults. Tribology International 2002, https://doi.org/10.1016/S0301-679X(02)00063-4.
39. Rafiee J, Rafiee M. A, Tse P. W. Application of mother wavelet functions for automatic gear and bearing fault diagnosis. Expert Systems with Applications 2010, https://doi.org/10.1016/j.eswa.2009.12.051.

40. Rai A, Upadhyay S. H. A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. Tribology International 2016; 289-306, https://doi.org/10.1016/j.triboint.2015.12.037.
41. Randall R. B, Antoni J. Rolling element bearing diagnostics-A tutorial. Mechanical Systems and Signal Processing 2011; 25(2): 485-520, https://doi.org/10.1016/j.ymssp.2010.07.017.
42. Ren L, Sun Y, Cui J, Zhang, L. Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. Journal of Manufacturing Systems 2018; 48: 71-77, https://doi.org/10.1016/j.jmsy.2018.04.008.
43. Rohlmann A, Schmidt H, Gast U, Kutzner I, Damm P, Bergmann G. In vivo measurements of the effect of whole body vibration on spinal loads. European Spine Journal 2014, https://doi.org/10.1007/s00586-013-3087-8.
44. Saha B, Goebel K, Christophersen J. Comparison of prognostic algorithms for estimating remaining useful life of batteries. Transactions of the Institute of Measurement and Contro 2009; 31(3-4): 293-308, https://doi.org/10.1177/0142331208092030.
45. Si X. S, Wang W, Hu C, Zhou D. Remaining useful life estimation - A review on the statistical data driven approaches. European Journal of Operational Research 2011; 213(1): 1-14, https://doi.org/10.1016/j.ejor.2010.11.018.
46. Sikorska J. Z, Hodkiewicz M, Ma L. Prognostic modelling options for remaining useful life estimation by industry. Mechanical Systems and Signal Processing 2011; 25: 1803-1836, https://doi.org/10.1016/j.ymssp.2010.11.018.
47. Soualhi A, Medjaher K. Zerhouni N. Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression. IEEE Transactions on Instrumentation and Measurement 2015; 64(1): 52-62, https://doi.org/10.1109/TIM.2014.2330494.
48. Souto Maior C. B, Moura M, Lins L. Droguett, Diniz H. E. Remaining Useful Life Estimation by Empirical Mode Decomposition and Support Vector Machine. IEEE Latin America Transactions 2016; 14(11): 4603-4610, https://doi.org/10.1109/TLA.2016.7795836.
49. Standardization. ISO 10816-7: Mechanical vibration - Evaluation of machine vibration by measurements on non-rotating parts. Part 7: Rotodynamic pumps for industrial applications, including measurements on rotating shafts. Switzerland: ISO. 2009.
50. Sutharssan T, Stoyanov S. Bailey C, Rosunally Y. Prognostics and health monitoring of high power LED. Micromachines 2012; 3: 78-100, https://doi.org/10.3390/mi3010078.
51. Sutrisno E, Oh H, Vasan A, Pecht M. Estimation of remaining useful life of ball bearings using data driven methodologies. 2012 IEEE Conference on Prognostics and Health Management 2012; 2: 1-7, https://doi.org/10.1109/ICPHM.2012.6299548.
52. Tandon N, Choudhury A. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. Tribology International 1999; 32(8): 469-480, https://doi.org/10.1016/S0301-679X(99)00077-8.
53. Torres M. E, Colominas M, Schlotthauer G, Flandrin P. A complete ensemble empirical mode decomposition with adaptive noise. IEEE International Conference on Acoustics, Speech and Signal Processing 2011, https://doi.org/10.1109/ICASSP.2011.5947265.
54. Vachtsevanos G, Lewis F, Roemer M, Hess A. Biqing Wu t al. Intelligent Fault Diagnosis and Prognosis for Engineering Systems. Intelligent Fault Diagnosis and Prognosis for Engineering Systems 2007, https://doi.org/10.1002/9780470117842.
55. Vapnik V. The Nature of Statistical Learning Theory. New York: Springer, 2000, https://doi.org/10.1007/978-1-4757-3264-1.
56. Wang L. Support Vector Machines : Theory and Applications. 2005, https://doi.org/10.1007/b95439.
57. Widodo A, Yang B. S. Machine health prognostics using survival probability and support vector machine. Expert Systems with Applications 2011; 38(7): 8430-8437, https://doi.org/10.1016/j.eswa.2011.01.038.
58. Wright S. J. Primal-Dual Interior-Point Methods. Primal-Dual Interior-Point Methods 2011.
59. Wu Z, Huang N. E. Ensemble Empirical Mode Decomposition: a Noise-Assisted Data Analysis Method. Advances in Adaptive Data Anal 2009; 1-41, https://doi.org/10.1142/S1793536909000047.
60. Yan R, Gao R, X. Chen X. Wavelets for fault diagnosis of rotary machines: A review with applications. Signal Processing 2014, 96(Part A): 1-15, https://doi.org/10.1016/j.sigpro.2013.04.015.

Caio Bezerra Souto MAIOR
Márcio das CHAGAS MOURA
Isis Didier LINS
Center for Risk Analysis and Environmental Modeling – CEERMA
Department of Production Engineering
Universidade Federal de Pernambuco – UFPE
Av. Prof. Moraes Rego
1235 – University City
Recife – PE – Brazil – 50670-901

Emails: caio.maior@ufpe.br, marcio@ceerma.org, isis.lins@ceerma.org

Article citation info:
BABISHIN V, TAGHIPOUR S. An algorithm for estimating the effect of maintenance on aggregated covariates with application to railway switch point machines. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 619–630, http://dx.doi.org/10.17531/ein.2019.4.11.

Vladimir BABISHIN
Sharareh TAGHIPOUR

AN ALGORITHM FOR ESTIMATING THE EFFECT OF MAINTENANCE ON AGGREGATED COVARIATES WITH APPLICATION TO RAILWAY SWITCH POINT MACHINES

ALGORYTM DO OCENY WPŁYWU KONSERWACJI NA ZAGREGOWANE ZMIENNE TOWARZYSZĄCEI JEGO ZASTOSOWANIE W ODNIESIENIU DO KOLEJOWYCH NAPĘDÓW ZWROTNICOWYCH

We propose an algorithm for estimating the effectiveness of maintenance on both age and health of a system. One of the main contributions is the concept of virtual health of the device. It is assumed that failures follow a nonhomogeneous Poisson process (NHPP) and covariates follow the proportional hazards model (PHM). In particular, the effect of maintenance on device's age is estimated using the Weibull hazard function, while the effect on device's health and covariates associated with condition-based monitoring (CBM) is estimated using the Cox hazard function. We show that the maintenance effect on the health indicator (HI) and the virtual HI can be expressed in terms of the Kalman filter concepts. The HI is calculated from Mahalanobis distance between the current and the baseline condition monitoring data. The effect of maintenance on both age and health is also estimated. The algorithm is applied to the case of railway point machines. Preventive and corrective types of maintenance are modelled as different maintenance effect parameters. Using condition monitoring data, the HI is calculated as a scaled Mahalanobis distance. We derive reliability and likelihood functions and find the least squares estimates (LSE) of all relevant parameters, maintenance effect estimates on time and HI, as well as the remaining useful life (RUL).

Keywords: *virtual health indicator, virtual age, maintenance effectiveness, preventive and corrective maintenance, Cox-Weibull hazard function, proportional hazards model.*

W artykule zaproponowano algorytm służący do szacowania skuteczności utrzymania ruchu w odniesieniu do wieku i stanu technicznego (kondycji) systemu. Główny wkład proponowanej metody stanowi koncepcja wirtualnego stanu urządzenia. Metoda zakłada, że uszkodzenia można zamodelować za pomocą niejednorodnego procesu Poissona, a zmienne towarzyszące za pomocą modelu proporcjonalnego hazardu. Mówiąc precyzyjniej, wpływ konserwacji na wiek urządzenia szacuje się z wykorzystaniem funkcji hazardu Weibulla, natomiast wpływ na stan urządzenia i zmienne towarzyszące związane z monitorowaniem stanu ocenia się stosując funkcję hazardu Coxa. W artykule pokazujemy, że wpływ konserwacji na wskaźnik stanu i wskaźnik stanu wirtualnego można wyrazić w kategoriach filtra Kalmana. Wskaźnik stanu oblicza się na podstawie odległości Mahalanobisa między bieżącymi a początkowymi danymi z monitorowania stanu. Ocenia się także wpływ utrzymania na wiek i kondycję systemu. Proponowany algorytm zastosowano w odniesieniu do napędów zwrotnicowych. Zapobiegawcze i naprawcze typy konserwacji zamodelowano jako różne parametry utrzymania ruchu. Korzystając z danych z monitorowania stanu, obliczono wskaźnik stanu jako skalowaną odległość Mahalanobisa. Wyprowadzono funkcje niezawodności i wiarygodności oraz obliczono metodą najmniejszych kwadratów szacunkowe wielkości wszystkich istotnych parametrów, a także szacunkowy wpływ konserwacji na wskaźniki czasu i stanu technicznego oraz pozostały okres użytkowania (RUL).

Słowa kluczowe: *wirtualny wskaźnik stanu technicznego, wiek wirtualny, skuteczność konserwacji, konserwacja zapobiegawcza i korygująca, funkcja hazardu Coxa–Weibulla, model proporcjonalnego hazardu.*

Notation

Typesetting Convention: vectors, matrices and arrays are indicated by arrows above the letters.

Latin Symbols

C	Cox model identifier.
CM	Corrective maintenance.
E	Expectation.
$\mathcal{L}$	Likelihood function.
L	Lubrication.
M	Maintenance.
PM	Preventive maintenance.
$\mathcal{R}$	Reliability.

TA	Thickness adjustment.
TS	Tightening of screws.
UCL	Upper confidence limit.
W	Weibull model identifier.
$\vec{X}$	Vector of covariates.
d	Mahalanobis distance.
i	Device index.
j	Manoeuvre index.
l	Number of observations in the baseline.
m	Number of covariates.
n	Total number of manoeuvres.
t	Time.
t_{Vj}	Virtual age after j th manoeuvre.
z	Health indicator.
z_{Vj}^{+}	Virtual health indicator after j th manoeuvre.
$*$	Optimality.

Greek Symbols

Λ	Cumulative hazard function.
α	Confidence level.
β	Shape parameter in Weibull distribution.
η	Scale parameter in Weibull distribution, characteristic life.
θ	Effect of maintenance on the health (as gauged by the health indicator) of the system.
λ	Power-law intensity (hazard) function.
φ	Effect of maintenance on the age of the system.
ω	Length of planning horizon (life cycle).

1. Introduction and background

Maintenance is critical for the longevity, reliability and availability of a vast majority of industrial, consumer and specialised systems and devices. However, a well-known postulate from reliability theory states that maintaining an entity (i.e. anything from the most basic component to a complex system) is justified and is beneficial only if the system displays a certain degradation in its performance with the passage of time. Such a deteriorating behaviour is called "aging", for the obvious analogy with the biological world. For this reason, in identifying the most effective maintenance, a common criterion for categorising maintenance actions is by effects these have on some general system metric, or parameter, which is usually age. In this regard, a common approach found in the literature on complex maintenance models of various industrial systems divides maintenance actions into four categories: worse repairs (increase the age when applied), minimal repairs (do not change the age when applied, leaving the system in the as-bad-as-old (ABAO) state), imperfect repairs (reduce the age by some factor between 0 and 1) and perfect repairs (effectively reduce the age to 0, amounting to as-good-as-new (AGAN) state) (Pulcini, 2003; Wu & Zuo, 2010). A preventive or corrective maintenance action affects the system's health state, and the effect of maintenance ranges from minimal (ABAO) to that equivalent to a complete renewal (AGAN). We are interested in measuring the maintenance effect and investigating how it impacts the system's health indicator (HI). The maintenance effect can range from 0 for AGAN state to 1 for ABAO state of the system.

Because the majority of real-life maintenance actions do not result in either ABAO or AGAN states, it is fair to state that, generally, maintenance actions amount to imperfect repairs (Pham & Wang, 1996), which may be classified into models featuring age reduction (Kijima & Nakagawa, Replacement policies of a shock model with imperfect maintenance, 1992), hazard rate reduction (Chan & Shaw, 1993), combined age-hazard reduction (Zhou, Xi, & Lee, 2007) and other models (Corman, Kraijema, Godjevac, & Lodewijks, 2017; Syamsundar, Muralidharan, & Naikan, General repair models for maintained systems, 2012). However, the age of a machine or even of a component is not always known. As an example, components or subsystems in protective devices, such as batteries in uninterruptible power supplies, may exhibit hidden failures, which are not manifested immediately, therefore making estimation of the age at failure difficult. Alternative methods for finding the optimal maintenance policy have been developed for different arrangements and systems subject to both evident and hidden failures, such as estimating the optimal number of minimal repairs before replacement (Babishin & Taghipour, Optimal maintenance policy for multicomponent systems with periodic and opportunistic inspections and preventive replacements, 2016; Babishin, Hajipoiur, & Taghipour, Optimisation of Non-Periodic Inspection and Maintenance for Multicomponent Systems, 2018; Babishin & Taghipour, Joint Maintenance and Inspection Optimization of a k-out-of-n System, 2016; Babishin & Taghipour, Joint Optimal Maintenance and Inspection for a k-out-of-n System, 2016).

Historically, imperfect repair has been quantified through improvement factors (Malik, 1979), (p, q) rule (Brown & Proschan, 1983), virtual age process (Uematsu & Nishida, 1987; Kijima, Some results for repairable systems with general repair, 1989) and superposed renewal process (Kallen, 2011), among others. Of those listed, the virtual age Models I and II due to Kijima assumed general repairs and utilised conditionally-distributed failure times (Kijima, Some results for repairable systems with general repair, 1989). Kijima's models were subsequently further developed by Dagpunar (Dagpunar, 1998), where functional dependency of the maintenance effect on both the time since previous maintenance action and the previous virtual age was assumed. Fuqing and Kumar (Fuqing & Kumar, 2012) generalised Kijima's Models I and II from constant to time-dependent repair effectiveness parameter (Fuqing & Kumar, 2012). Using Kijima's modelling framework, Doyen and Gaudoin classify the effects of maintenance as having a failure intensity-reducing, or an age-reducing effect, also allowing for a Markovian memory property (Doyen & Gaudoin, Classes of imperfect repair models based on reduction of failure intensity or virtual age, 2004). Furthering the framework of Kijima (Kijima, Some results for repairable systems with general repair, 1989) and Doyen and Gaudoin (Doyen & Gaudoin, Classes of imperfect repair models based on reduction of failure intensity or virtual age, 2004), in the present paper, virtual age and virtual health indicator are used, and the effects of maintenance are considered simultaneously on both intensity and age.

Maintenance optimisation in railway-related applications is considered, for example, by Corman et al. (Corman, Kraijema, Godjevac, & Lodewijks, 2017), where they propose a data-driven approach to optimising preventive maintenance of a light rail braking system in terms of reliability, availability and maintenance cost. Based on the data, they model reliability degradation by a Weibull distribution and use sequential optimisation to find optimal preventive maintenance intervals resulting in 30 % cost reduction (Corman, Kraijema, Godjevac, & Lodewijks, 2017). Corman et al. further suggest using multi-component optimisation to capture complex economic and structural dependence (Corman, Kraijema, Godjevac, & Lodewijks, 2017).

In the context of many repairable systems, "events" can be considered points at which a system changes its state, or exchanges information with its surroundings. Common events include failures, inspections and various kinds of maintenance. Identifying these properly and unambiguously, however, can be challenging, if the effects of such events are not readily observable.

An aspect of interest to the present investigation is the type of maintenance, classified into preventive maintenance (PM) and corrective maintenance (CM). Doyen and Gaudoin proposed a model for each type of PM and CM, each with just one maintenance policy available (Doyen & Gaudoin, Imperfect maintenance in a generalized competing risks framework, 2006). Nasr *et al.* consider failure-point virtual age for CM and repair-point virtual age for PM (Nasr, Gasmi, & Sayadi, 2013). Said and Taghipour further expanded this by considering three maintenance types for PM events and minimal repair for CM events (Said & Taghipour, 2016). They derive the likelihood function for estimating the parameters of the failure process and the effects of preventive maintenance, as well as provide the conditional reliability and the expected number of failures between two consecutive PM types (Said & Taghipour, 2016). Other methods included using feed-forward artificial neural networks (ANN) on condition monitoring data with asset targets' being asset survival probabilities estimated by Kaplan-Meier (KM) and degradation failure probability density function (PDF) estimator (Heng, et al., 2009).

Reliability and availability of multicomponent systems were obtained, for example, in (Babishin & Taghipour, Optimal maintenance policy for multicomponent systems with periodic and opportunistic inspections and preventive replacements, 2016; Babishin, Hajipoiur, & Taghipour, Optimisation of Non-Periodic Inspection and Maintenance for Multicomponent Systems, 2018). Chen *et al.* use queueing theory to find reliability and availability expressions for a 2-component cold standby system with repairman who may have vacation under Poisson shocks (Chen, Meng, & Chen, 2014). For more complex systems, however, Monte Carlo simulation is widely used, such as in Wang and Cotofana (Wang & Cotofana, 2010), Conn *et al.* (Conn, Deleris, Hosking, & Thorstensen, 2010) and Lim and Lie (Lim & Lie, 2000). Bayesian methods have also been used to estimate the parameters for reliability and maintainability in Nasr *et al.* (Nasr, Gasmi, & Sayadi, 2013), Yu *et al.* (Yu, Song, & Cassady, 2008) and Fuqing and Kumar (Fuqing & Kumar, 2012). In addition, Nasr *et al.* (Nasr, Gasmi, & Sayadi, 2013) derive log-likelihood functions corresponding to failure-point and repair-point virtual age models (Nasr, Gasmi, & Sayadi, 2013). In this paper, both reliability and log-likelihood expressions are provided.

Presently, a large-scale move towards Internet of Things (IoT) is being implemented in various industries. This makes the data from monitoring equipment and sensors ever more ubiquitous and accessible. With this in mind, a question arises as to how to incorporate such operating condition data into the reliability models. One widely-used method is to treat condition monitoring or operating condition data as covariates within the Cox proportional hazard models' framework (Syamsundar & Naikan, Imperfect repair proportional intensity models for maintained systems, 2011; Cox, 1992; Bendell, Wightman, & Walker, 1991). An obstacle to the universality of such models is that they assume that covariates are time-independent, thus ignoring any influence of changing operating conditions. Previously, accelerated failure time model (AFTM) has been incorporated with virtual age model by Martorell *et al.* (Martorell, Sanchez, & Serradell, 1999). However, combining imperfect repair models with either proportional hazards model or AFTM and considering the effect of covariates is rare, and the attempts found in the literature adopt some simplifying assumptions, such as piecewise-constant operating conditions (Hu, Jiang, & Liao, 2017). Proportional hazards model has also been applied to covariate data for railway maintenance effectiveness estimation in (Babishin & Taghipour, Maintenance Effectiveness Estimation with Applications to Railway Industry, 2019).

Cha and Finkelstein (Cha & Finkelstein, 2016) considered periodic and age-based imperfect PM and minimal repairs in-between (Cha & Finkelstein, 2016). In the present paper, however, neither PM, nor CM events are limited to minimal or perfect repairs, which makes the model more general and widely applicable.

Predicting degradation of a system, machine or device and choosing the best maintenance actions allow preventing or reducing its damage or failure. This is where prognostics and health management (PHM) becomes important. We make use of condition monitoring data, which are observations of different parameters (e.g. temperature, weather, current, voltage). Galar et al. previously proposed feature extraction through data reduction, where only significant data are retained, and irrelevant information is discarded (Galar, Gustafson, Tormos, & Berges, 2012). These observations are aggregated into a health indicator, which represents the system's condition. Health indicator was used by Kumar et al. for detecting the degradation of electronic products (personal computers) (Kumar, Vichare, Dolev, & Pecht, 2012). Their health indicator represents a weighted sum of the fractional contributions of each bin in a time window (Kumar, Vichare, Dolev, & Pecht, 2012).

In repairable systems, the passage of time, the number of operating cycles and/or the changes in the system's operating conditions signify deterioration of the system and its approaching failure. This motivates preventive maintenance, which improves the system's condition and extends its remaining useful life (RUL). RUL is defined as "the expected number of remaining manoeuvres that can be achieved before reaching the failure state" (Letot, et al., 2015).

The main objectives of the present research are to demonstrate an algorithm for quantifying the effectiveness of corrective and preventive maintenance performed on a machine, and to estimate the machine's degradation rate and remaining useful life, given the maintenance effectiveness.

In the current paper, condition monitoring data are used for estimating the effect of maintenance on both the age of a railway point machine and its covariates. A railway switch, or point machine, is a device for allowing the trains to pass from one railway track onto another one, which makes these devices both necessary and ubiquitous for simultaneous operation of trains in multiple directions. A manoeuvre is a 7-phase sequence of operations performed by components of a point machine (Letot, et al., 2015).

Because of the function point machines perform, they greatly affect the service of rail transportation. This, in turn, affects the safety of passengers, the economic benefits, efficiency and timeliness of train travel. All of these factors can potentially incur huge costs and penalties, including loss of life from accidents, if the system does not perform as expected. For this reason, excessive funds are spent every year on inspection and maintenance of such systems as point machines in order to minimise their failures and to ensure they perform correctly and reliably. For example, the Swedish Rail Administration estimates the costs of railway track maintenance falling under the category of switches and crossings to account for almost 1/3 of the total maintenance costs (Innotrack, 2009). Thus, improving reliability and maintainability in this sector may not only result in the improved

safety and lower accident occurrence, but can also bring significant cost reductions to the railroad industry.

Condition monitoring and health management of railway assets, such as point machines, has received some coverage in the literature (Atamuradov, Medjaher, Dersin, Lamoureux, & Zerhouni, 2017; Ardakani, et al., 2012). For example, Ardakani et al. (Ardakani, et al., 2012) use feature extraction techniques and principal component analysis (PCA) as the methods for prognostics and health management for analysing the degradation of electromechanical point machines for railway turnouts. A turnout is a point machine with the switch rails connected to it.

The present article is structured as follows: Section 2 contains the relevant background; Section 3 presents the model; Section 4 contains reliability and likelihood functions; Section 5 illustrates the models by providing numerical examples; lastly, Section 6 summarises the conclusions.

2. Model

2.1. Health indicator calculation

Since a maintenance action can affect the age of a system as well as the condition monitoring data, we investigate both effects. More specifically, we estimate how much reduction in the system's age is caused by a maintenance type, and how the health indicator (which is constructed solely based on the condition monitoring data) is affected by the maintenance action. Health indicator is a measure quantifying the deterioration of the system.

At each operational actuation of the machine, readings from the sensors and diagnostic modules monitoring such parameters, as temperature, humidity, voltage, current, etc. are recorded. Each of the monitoring parameters is designated an index m (e.g. for temperature, $m=1$, for humidity $m=2$, etc.). The ordinal number of an actuation is designated as j and used as a counting index (e.g. for the 2000th actuation of a point machine, j=2000). These are then aggregated to form covariate X_{m_j}. The health indicator, denoted as z_j, is obtained from Mahalanobis distance (MD) calculation as follows:

$$z_j = \frac{\sqrt{\left(\vec{X}_j - \vec{\mu}\right)^T \left[\operatorname{cov}\left(\vec{X}_j\right)\right]^{-1} \left(\vec{X}_j - \vec{\mu}\right)}}{\left(\chi^2\right)^{-1}(0.999999, m)} =$$

$$= \frac{\sqrt{\left[X_{1_j} - \mu_{1_l} \quad \cdots \quad X_{m_j} - \mu_{m_l}\right]\left[\operatorname{cov}\left(\vec{X}_j\right)\right]^{-1} \begin{bmatrix} X_{1_j} - \mu_{1_l} \\ \vdots \\ X_{m_j} - \mu_{m_l} \end{bmatrix}}}{\left(\chi^2\right)^{-1}(0.999999, m)} \tag{1}$$

where j denotes the number of actuations, $\vec{X}_j$ is the vector of m covariates for j th actuation, $\vec{X}_j = \left[X_{1_j} \quad \cdots \quad X_{m_j}\right]$, where $j = 1, 2, \ldots, n$, $(\chi^2)^{-1}(0.999999, m)$ is the value of inverse cumulative distribution of the 0.999999th quantile of a chi-squared distribution with m degrees of freedom, which denotes the threshold for the "healthy" values of the HI, $\vec{\mu}$ is the vector of means over l observations, also called "baseline", such that:

$$\vec{\mu} = \left[\mu_{1_l} \quad \cdots \quad \mu_{m_l}\right]$$

$$= \left[\sum_{l=1}^{m} X_{1_l} / m \quad \cdots \quad \sum_{l=1}^{m} X_{m_l} / m\right], \tag{2}$$

Also, $\left[\operatorname{cov}\left(\vec{X}_j\right)\right]^{-1}$ is the inverse of the covariate matrix $\left[\operatorname{cov}\left(\vec{X}_j\right)\right]$, given as:

$$\left[\operatorname{cov}\left(\vec{X}_j\right)\right] =$$

$$= \begin{bmatrix} \operatorname{var}\left(X_{1_j}\right) & \operatorname{cov}\left(X_{1_j}, X_{2_j}\right) & \cdots & \operatorname{cov}\left(X_{1_j}, X_{m_j}\right) \\ \operatorname{cov}\left(X_{2_j}, X_{1_j}\right) & \operatorname{var}\left(X_{2_j}\right) & \cdots & \operatorname{cov}\left(X_{2_j}, X_{m_j}\right) \\ \vdots & \cdots & \ddots & \vdots \\ \operatorname{cov}\left(X_{m_j}, X_{1_j}\right) & \operatorname{cov}\left(X_{m_j}, X_{2_j}\right) & \cdots & \operatorname{var}\left(X_{m_j}\right) \end{bmatrix}. \tag{3}$$

Thus, when HI < 1, the MD is considered to be chi-squared distributed, and the system is "healthy". When HI $\geq$ 1, the probability that the covariates are normally distributed and their covariances are chi-squared distributed is very small, which suggests that the system is demonstrating "abnormal" behaviour.

In general, the extent to which a machine has moved away from its "baseline", or usual operation, is quantified by the HI. The expectation here is that a large deviation from baseline signals an ongoing degradation of the system and, as a result, increases failure risk. When the health indicator is below the predetermined threshold ($HI < 1$), the system is operating normally. Consequently, defining the alternative event, have $HI \geq 1$, which corresponds to the "failed" operational state of the system.

2.2. Virtual health and the effect of maintenance on the system's health indicator ("Cox model")

When the ratio of the hazards for different treatments does not change with time, proportional hazards models can be used to describe the reliability of the system.

2.2.1. Virtual health indicator algorithm

We consider failures as having a negative effect on the HI. The effect of failures on the HI is modelled using a Cox proportional hazards model, where the hazard function λ_C is given for each machine as:

$$\lambda_C\left(z, \theta_M\right) = \exp\left\{\theta_M \gamma z\right\}, \tag{4}$$

where z is the HI of the machine, γ is the Cox regression coefficient used for scaling the covariates and θ_M is the maintenance effect on machine's HI.

In order to capture the effects of each maintenance type and isolate them from the cumulative effects of maintenance events which have taken place in the past history, the health indicator values (Mahalanobis distances) have to be scaled by the maintenance effect factor (MEF) θ_M after the maintenance events. The virtual health indicator is denoted as $z_{V_j}^{+}$, with "V" standing for "virtual" and "+" indicating that it is recalculated after a maintenance event has taken place in order to account for the effect of the most recent maintenance.

The procedure to calculate the maintenance effect factor is as follows.

Given:

$z_i \geq 0, z_0 = 0, z_{V\,j}^- = z_j^-, z_{V1}^+ = z_1^+,\ 0 \leq \theta_{PM} \leq 1,\ 0 \leq \theta_{CM} \leq 1.$

Obtain:

1. Take the HI before the first maintenance event to be z_1^- and after it to be z_1^+.
2. Calculate the first maintenance effectiveness using the following expression:

$$\theta_{M_1} = \frac{z_{V1}^+}{z_1^-} = \frac{z_1^+}{z_1^-} \tag{5}$$

3. Take the HI after the first maintenance and just prior to the second maintenance to be z_2^-.
4. Taking the HI just after the second maintenance z_2^+ from the data for the manoeuvre immediately following the second maintenance event, calculate preliminary estimate of maintenance effect $\hat{\theta}_{M_2}$ as:

$$\hat{\theta}_{M_2} = \frac{z_2^+}{z_2^-} \tag{6}$$

5. Estimate the value of the virtual HI z_{V2}^+ after the second maintenance event using the following formula:

$$z_{V2}^+ = \left(z_2^- - z_1^+\right)\hat{\theta}_{M_2} + z_1^- \theta_{M_1} = \left(z_2^- - z_1^+\right)\hat{\theta}_{M_2} + z_1^+ \tag{7}$$

6. Calculate the new estimate of maintenance effectiveness θ_{M2} using the virtual health indicator as follows:

$$\theta_{M2} = \frac{z_{V2}^+}{z_2^-}. \tag{8}$$

7. Repeat the steps above to calculate new maintenance effectiveness estimates for events 3,4,…,*i* by induction using the following recursive formula for step 5:

$$\begin{aligned} z_{V_j}^+ &= \left(z_j^- - z_{V_{j-1}}^+\right)\hat{\theta}_{M_j} + z_{V_{j-1}}^+ = \\ &= \left(z_j^- - z_{V_{j-1}}^+\right)\hat{\theta}_{M_j} + \left(z_{j-1}^- - z_{V_{j-2}}^+\right)\hat{\theta}_{M_{j-1}} + \cdots + \left(z_2^- - z_{V_2}^+\right)\hat{\theta}_{M_2} + z_1^-\hat{\theta}_{M_1} = \\ &= \left(z_j^- - z_{V_{j-1}}^+\right)\frac{z_j^+}{z_j^-} + \left(z_{j-1}^- - z_{V_{j-2}}^+\right)\frac{z_{j-1}^+}{z_{j-1}^-} + \cdots + \left(z_2^- - z_{V_2}^+\right)\frac{z_2^+}{z_2^-} + z_1^+. \end{aligned} \tag{9}$$

For step 6 of the current procedure, use the following formula:

$$\theta_{M_j} = \frac{z_{V\,j}^+}{z_j^-} \tag{10}$$

In order to better visualize the calculation procedure and the formulae, Figure 1 below represents a general case of a deteriorating machine or device subject to imperfect maintenance. In such case, the first maintenance action (denoted as M_1) will result in the virtual HI closest to the baseline, thus representing the largest health-improving effect, followed by the virtual HI for the second maintenance M_2 and so on. Note that the horizontal axis in the figure represents the distance from the baseline (or 0), and not the time progression. In Figure 1, the segment [z_0 ; z_1^+] represents the virtual health $\theta_{M_1} z_1^-$ of the device after the first maintenance action has been performed (i.e.

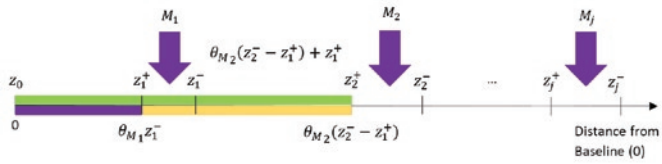


Fig. 1. Visualisation of maintenance events and procedure for estimating their effects

the distance from the baseline to the manoeuvre right after the first maintenance event). The segment [z_1^+ ; z_2^+] represents deterioration of the virtual health $\theta_{M_2}\left(z_2^- - z_1^+\right)$, occurring between the first and the second maintenance events and calculated right after the second maintenance event. The segment [z_0 ; z_2^+] equal in length to the combined segments [z_0 ; z_1^+] and [z_1^+ ; z_2^+] represents the virtual health $\theta_{M2}\left(z_2^- - z_1^+\right) + z_1^+$ after the second maintenance.

The virtual HI is calculated for each machine using θ_{PM} and θ_{CM} to denote the effect of, respectively, preventive and corrective maintenance on the former as follows:

$$\begin{aligned} j = 1:\quad & z_{V_1}^+ = \theta_{M_1} z_1^-, \\ j = 2:\quad & z_{V_2}^+ = \theta_{M_2} z_{V_2}^- = \theta_{M_2}\left(z_{V_1}^+ + (z_2^- - z_1^+)\right) = \theta_{M_2}\left(\theta_{M_1} z_1^- + (z_2^- - z_1^+)\right), \\ j = 3:\quad & z_{V_3}^+ = \theta_{M_3}\left(z_{V_2}^+ + (z_3^- - z_2^+)\right) = \theta_{M_3}\left(\theta_{M_2}\left(\theta_{M_1} z_1^- + (z_2^- - z_1^+)\right) + (z_3^- - z_2^+)\right), \\ & \vdots \\ j = n:\quad & z_{V_n}^+ = \theta_{M_n} z_{V_n}^- = \theta_{M_n}\Big(\theta_{M_{n-1}}\big(\ldots\left(\theta_{M_2}\left(\theta_{M_1} z_1^- + (z_2^- - z_1^+)\right) + (z_3^- - z_2^+)\right) + \cdots\big) \\ & \quad + (z_n^- - z_{n-1}^+)\Big), \end{aligned}$$

$$\theta_{M_j} = \begin{cases} \theta_{PM_j}, \text{if maintenance event } j \text{ is a PM;} \\ \theta_{CM_j}, \text{if maintenance event } j \text{ is a CM.} \end{cases} \tag{11}$$

where z_j^- is the value of HI calculated right before the maintenance action, z_j^+ is the value of HI calculated right after the maintenance action, and superscript M denotes the type of maintenance action.

It can be noted from Eq. 9 that the form of the virtual health indicator estimate is identical to the current state estimate of a Kalman filter [21, 26]:

$$z_{V_j}^+ = \left(z_j^- - z_{V_{j-1}}^+\right)\hat{\theta}_{M_j} + z_{V_{j-1}}^+ \quad \text{cf.} \quad Est_t = (Meas - Est_{t-1})K_G + Est_{t-1}, \tag{12}$$

where Est_t is the current estimate of the state, $Meas$ is the initial measurement, Est_{t-1} is the initial estimate of the state, K_G is the Kalman gain, and so have:

$$z_{V_j}^+ = Est_t, \quad z_j^- = Meas, \quad z_{V_{j-1}}^+ = Est_{t-1}, \quad \hat{\theta}_{M_j} = K_G. \tag{13}$$

Furthermore, maintenance effectiveness estimate $\hat{\theta}_{M_j}$ can be compared to Kalman filter gain using Eq. 6 and identities from Eq. 13, so that:

$$\hat{\theta}_{M_j} = \frac{z_j^+}{z_j^-} = \frac{z_j^+}{Meas} \quad \text{cf.} \quad K_G = \frac{Er_{Est_t}}{Er_{Est_t} + Er_{Meas_t}}, \tag{14}$$

and from $\hat{\theta}_{M_j}$=K_G (Eq. 13) and Eq. 14 it follows that:

$$K_G = \frac{z_j^+}{Meas},$$

$$z_j^+ = K_G \cdot Meas = \frac{Meas \cdot Er_{Est_t}}{Er_{Meas_t} + Er_{Est_t}} \tag{15}$$

where Er_{Est_t} is the error in the estimate of the state and Er_{Meas_t} is the error in the measurement of the state. Thus, the health indicator after a failure or maintenance event can be interpreted using Kalman filter theory as the initial measurement of the state multiplied by Kalman gain. It can also be expressed through the initial measurement of the state multiplied by the error in the current estimate and divided by the total error of the initial measurement and that of the current estimate.

In addition, from Eq. 10 and Eq. 13 have:

$$\theta_{M\,j} = \frac{Est_t}{Meas} \tag{16}$$

Analysing the formulae for the calculation of the maintenance effect factors θ_{PM} and θ_{CM}, it can be seen that:

$$\begin{aligned} &\theta_{M\,j} < 0 \text{ if and only if either :} \\ &z_{V\,j}^+ < z_j^+ < z_j^-, \text{ or} \\ &z_j^- < z_{V\,j-1}^+ < z_{V\,j}^+ . \end{aligned} \tag{17}$$

Similarly, rewriting Eq. 17 using Kalman filter notation:

$$\begin{aligned} &\theta_{M\,j} < 0 \text{ if and only if either :} \\ &Est_t < \frac{Meas \cdot Er_{Est_t}}{Er_{Meas_t} + Er_{Est_t}} < Meas, \text{ or} \\ &Meas < Est_{t-1} < Est_t . \end{aligned} \tag{18}$$

Both Eq. 17 and Eq. 18 describe cases in which the system experiences improvement in HI as it ages and which violate the basic characteristics of repairable systems. Thus, $\theta_{M\,j} < 0$ can serve as an indicator that the system experiences "early mortality" and its hazard function is decreasing with the system's age.

2.3. Virtual age and the effect of maintenance on the system's age ("Weibull model")

Whenever a system is subject to degradation with time, the latter is commonly modelled as affecting the system's age. In the context of the present problem, it is assumed that each machine is subject to a nonhomogeneous Poisson process (NHPP) with the time-dependent power law intensity function λ_W of the general form:

$$\lambda_W(t, \varphi_M) = \frac{\beta}{\eta}\left(\frac{\varphi_M t}{\eta}\right)^{\beta-1}, \tag{19}$$

where β is the Weibull shape parameter, η is the Weibull scale parameter, t is the time to failure, φ_M is the maintenance effect on system's age and $M = \begin{cases} PM, if\ preventive\ maintenance\ is\ performed \\ CM, if\ corrective\ maintenance\ is\ performed \end{cases}$

Assuming that the effect of maintenance on age is cumulative, it is modelled through the concept of virtual age.

2.3.1. Virtual age

Using φ_{PM} and φ_{CM} to denote the effect of, respectively, preventive and corrective maintenance on machine's age, so that $0 \le \varphi_{PM} \le 1$, $0 \le \varphi_{CM} \le 1$, where 0 corresponds to the as-good-as-new (AGAN) state and 1 to the as-bad-as-old (ABAO) state, and designating virtual age for the $_j$th aintenance action as t_{V_j}, obtain:

$$j = 1: \begin{cases} t_{V_1}^{PM} = \varphi_{PM}(t_1 - t_0), \text{if current event is a PM;} \\ t_{V_1}^{CM} = \varphi_{CM}(t_1 - t_0), \text{if current event is a CM;} \end{cases}$$

$$j = 2: \begin{cases} t_{V_2}^{PM} = \varphi_{PM}\left(t_2 - t_1 + t_{V_1}^{PM}\right), \text{if current event is} \\ \text{a PM and previous event was a PM;} \\ t_{V_2}^{PM} = \varphi_{PM}(t_2 - t_1), \text{ if current event is} \\ \text{a PM and previous event was a CM;} \\ t_{V_2}^{CM} = \varphi_{CM}\left(t_2 - t_1 + t_{V_1}^{PM}\right), \text{ if current event is} \\ \text{a CM and previous event was a PM;} \\ t_{V_2}^{CM} = \varphi_{CM}(t_2 - t_1), \text{ if current event is} \\ \text{a CM and previous event was a CM;} \end{cases} \tag{20}$$

$$j = n: \begin{cases} t_{V_n}^{PM} = \varphi_{PM}\left(t_n - t_{n-1} + t_{V_{n-1}}^{PM}\right), \text{if current event is a} \\ \text{PM and previous event was a PM;} \\ t_n^{PM} = \varphi_{PM}(t_n - t_{n-1}), \text{if current event is a} \\ \text{PM and previous event was a CM;} \\ t_{V_n}^{CM} = \varphi_{CM}\left(t_{n,} - t_{n-1} + t_{V_{n-1}}^{PM}\right), \text{if current event is a} \\ \text{CM and previous event was a PM;} \\ t_n^{CM} = \varphi_{CM}(t_n - t_{n-1}), \text{if current event is a} \\ \text{CM and previous event was a CM;} \end{cases}$$

It can be noted that the value of 0 for the effect of maintenance on the age indicates a complete renewal of the system, and the value of 1 is analogous to the minimal repair.

In the present subsection, a Weibull model for an NHPP failure process has been discussed for identifying the effect of a particular maintenance type on the age of a component or a device. The available condition monitoring data are incorporated into maintenance decision-making trough the Cox proportional hazards model. This is a useful technique for estimating reliability and related metrics.

2.4. Combined (Cox-Weibull) model

Point machines have subassemblies and components that experience age-dependent deterioration (e.g. gearbox) and those that do not (e.g. electronic control and diagnostic module). Thus, the importance of condition-based vs. age-based maintenance estimation techniques depends on the particular component. Moreover, modern monitoring and diagnostic capabilities within the IoT framework provide plenty of condition monitoring data in addition to the age-based data.

In the preceding subsections, two models were discussed: a Cox PHM model, which quantifies the effect of maintenance on the health indicator, and a Weibull model, which identified the effect of a particular maintenance type on age. Thus, in estimating the hazard function for a point

machine as a whole, the available data can be taken into consideration by combining the age-based hazard in the form of Weibull hazard function with the condition-based monitoring hazard in the form of Cox proportional hazards model. In the present section, these models are combined to obtain a more powerful model.

In order to improve the sensitivity and applicability of the model, the Cox-Weibull model was enhanced with the maintenance effectiveness estimates multiplicative to the virtual age and virtual health indicator. The model allows to reset the health indicator to the value reflecting the maintenance effectiveness and the system's state by multiplying the health indicator after the specific type of maintenance by the maintenance effect factor for that particular maintenance type. The visualization of the model is given in Figure 2 below.

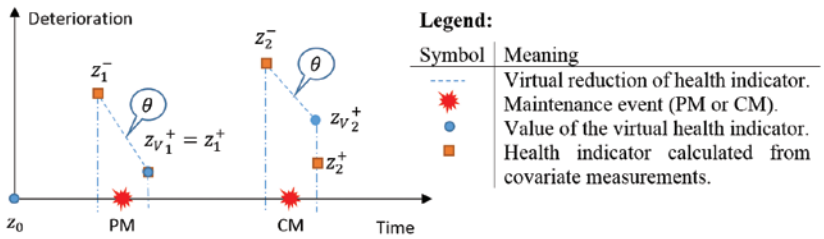


Fig. 2. Visualisation of two sample maintenance events with virtual health indicator

In Figure 2, squares indicate points at which condition monitoring data, or covariates are recorded just before and after a system event (such as failure, or maintenance). Circles represent points at which virtual health indicator is calculated. Following the performance of preventive maintenance (PM) (indicated by an oval callout with θ inside), the device's health is improved and its deterioration is reduced. This reduction is reflected in the changes within the condition monitoring and/or covariate data, which results in a decrease of HI as shown by the square markers. With the use of the device and the passage of time, it keeps deteriorating to failure. At this point, corrective maintenance (CM) is performed, HI is reduced and the device's health is improved. While HI shows a large improvement as represented by square markers, it is not clear how much of a contribution did the most recent maintenance action have compared to the previous maintenance history. Such a reduction in HI is likely due to the cumulative effect of all the previous maintenance actions. However, of interest is the isolated effect of each maintenance type, such as PM and CM, since these most likely happened intermittently in the past operational history.

With this goal, the previously-presented Weibull and Cox models are combined together to improve the sensitivity of the model and to quantify the effects of PM and CM maintenance types on the age and health of the device or system. The hazard function $\lambda(t,z,\varphi_M,\theta_M)$ for the new combined Cox-Weibull model has the following form:

$$\lambda(t,z,\varphi_M,\theta_M)=\frac{\beta}{\eta}\left(\frac{\varphi_M t}{\eta}\right)^{\beta-1}\exp\{\theta_M\gamma z\}=\frac{\beta}{\eta}\left(\frac{t_V^M}{\eta}\right)^{\beta-1}\exp\left\{\gamma z_{V_j}^+\right\} \tag{21}$$

where all the terms are as previously described.

The cumulative hazard function is then given as follows:

$$\Lambda\left(t_{V_j}^M\right)=\int_0^t\lambda(t,z,\varphi_M,\theta_M)=\frac{\theta_M\gamma z}{\varphi_M}\left(\frac{\varphi_M t_{V_j}^M}{\eta}\right)^{\beta}. \tag{22}$$

In order to establish the dynamics of the hazard function and to infer whether its form is suitable for a particular case at hand, we take the derivative of $\lambda(t,z,\varphi_M,\theta_M)$ with respect to time as follows:

$$\lambda'(t,z,\varphi_M,\theta_M)=\frac{d}{dt}\left[\frac{\beta}{\eta}\left(\frac{\varphi_M t}{\eta}\right)^{\beta-1}\exp\{\theta_M\gamma z\}\right]=\frac{\beta(\beta-1)\varphi_M{}^{\beta-1}t^{\beta-2}}{\eta^{\beta}}\exp\{\theta_M\gamma z\} \tag{23}$$

It should be noted that both maintenance effect indicators φ_M,θ_M satisfy the Markovian property, since they depend only on the preceding state and not the entire evolution of the states up to the present. Thus, they can be treated as time-independent.

Setting the derivative of the hazard function equal to 0, we can find the critical points:

$$\lambda'(t,z,\varphi_M,\theta_M)=0$$
$$\frac{\beta(\beta-1)\varphi_M{}^{\beta-1}t^{\beta-2}}{\eta^{\beta}}\exp\{\theta_M\gamma z\}=0. \tag{24}$$

Solving Eq. 24, obtain different cases:

$$\begin{cases}\beta=1:\lambda=\text{const.}\\ \beta=0:\lambda=0\\ \varphi_M=0:\text{purely AGAN maintenance effect}\end{cases} \tag{25}$$

In the case of $\lambda=$ const., failure distribution is an exponential distribution, and there is no benefit from performing any maintenance activities, since failures result not from deterioration, but rather from random events. In the case of $\lambda=0$, the entire hazard function is 0, and the system is not deteriorating. In the case of $\varphi_M=0$, each maintenance is perfect and results in as-good-as-new state, thus being equivalent in effect to replacement.

Using the hazard and cumulative hazard functions as given in Eq. 21 and Eq. 22, reliability and likelihood functions are constructed in order to estimate the optimal parameters of interest.

3. Reliability and likelihood functions

The goal of the present methodology is to estimate simultaneously the parameters β and η of the power law intensity function, as well as the maintenance effectiveness estimates φ_{PM}, φ_{CM}, θ^{PM}, θ^{CM}, and the coefficients of the covariates γ_i. All of these can be aggregated into a vector $\vec{p}$:

$$\vec{p}=(\beta,\eta,\varphi_{PM},\varphi_{CM},\theta_{PM},\theta_{CM},\gamma). \tag{26}$$

First, the reliability function is calculated by taking into account the suspension histories due to preventive maintenance, as well as failures and pseudo failures (i.e. when the health indicator crosses some threshold). Then, the likelihood function of the model is calculated.

3.1. Reliability

Different cases require different reliability function calculations, as shown below. All of the expressions are given for each device i.

Case 1: event j is a failure, immediately followed by CM:

- Previous event (j-1) is a failure, followed by CM:

$$f\left(t_{V_j}^{CM}-t_{V_{j-1}}^{CM}\middle|t_{V_{j-1}}^{CM},z_{V_{j-1}}^+\right)= \tag{27}$$
$$=\lambda\left(t_{V_j}^{CM}\right)\exp\left\{-\left(\Lambda\left(t_{V_j}^{CM}\right)-\Lambda\left(t_{V_{j-1}}^{CM}\right)\right)\right\}.$$

• Previous event $(j-1)$ is a PM:

$$f\left(t_{V_j}^{CM}-t_{V_{j-1}}^{PM}\middle|t_{V_{j-1}}^{PM},z_{V_{j-1}}^{+}\right)=\lambda\left(t_{V_j}^{CM}\right)\exp\left\{-\left(\Lambda\left(t_{V_j}^{CM}\right)-\Lambda\left(t_{V_{j-1}}^{PM}\right)\right)\right\}$$

$$=\frac{\beta}{\eta}\left(\frac{\varphi_{CM}t_{V_j}^{CM}}{\eta}\right)^{\beta-1}\exp\left\{\theta_{CM}\gamma_i z_i-\left(\frac{\theta^{CM}\gamma_i z_i}{\varphi_{CM}}\left(\frac{\varphi_{CM}t_{V_j}^{CM}}{\eta}\right)^{\beta}-\right.\right.$$

$$\left.\left.-\frac{\theta^{PM}\gamma_i z_i}{\varphi_{PM}}\left(\frac{\varphi_{PM}t_{V_{j-1}}^{PM}}{\eta}\right)^{\beta}\right)\right\}=\frac{\beta}{\eta}\left(\frac{\varphi_{CM}t_{V_j}^{CM}}{\eta}\right)^{\beta-1}\exp\left\{\frac{\gamma_i z_i}{\varphi_{CM}}\left(\theta^{CM}\varphi_{CM}-\right.\right.$$

$$\left.\left.-\theta^{CM}\left(\frac{\varphi_{CM}t_{V_j}^{CM}}{\eta}\right)^{\beta}+\theta^{PM}\left(\frac{\varphi_{PM}t_{V_{j-1}}^{PM}}{\eta}\right)^{\beta}\right)\right\}. \quad (28)$$

Case 2: event j is a PM:

• Previous event $(j-1)$ is a failure, followed by CM:

$$R\left(t_{V_j}^{PM}-t_{V_{j-1}}^{CM}\middle|t_{V_{j-1}}^{CM},z_{V_{j-1}}^{+}\right)=\exp\left\{-\left(\Lambda\left(t_{V_j}^{PM}\right)-\Lambda\left(t_{V_{j-1}}^{CM}\right)\right)\right\}$$

$$=\exp\left\{-\gamma_i z_i\left(\frac{\theta_{PM}}{\varphi_{PM}}\left(\frac{\varphi_{PM}t_{V_j}^{PM}}{\eta}\right)^{\beta}-\frac{\theta_{CM}}{\varphi_{CM}}\left(\frac{\varphi_{CM}t_{V_{j-1}}^{CM}}{\eta}\right)^{\beta}\right)\right\}. \quad (29)$$

• Previous event $(j-1)$ is a PM:

$$R\left(t_{V_j}^{PM}-t_{V_{j-1}}^{PM}\middle|t_{V_{j-1}}^{PM},z_{V_{j-1}}^{+}\right)=\exp\left\{-\left(\Lambda\left(t_{V_j}^{PM}\right)-\Lambda\left(t_{V_{j-1}}^{PM}\right)\right)\right\}$$

$$=\exp\left\{-\gamma_i z_i\left(\frac{\theta_{PM}}{\varphi_{PM}}\left(\frac{\varphi_{PM}t_{V_j}^{PM}}{\eta}\right)^{\beta}-\frac{\theta_{PM}}{\varphi_{PM}}\left(\frac{\varphi_{PM}t_{V_{j-1}}^{PM}}{\eta}\right)^{\beta}\right)\right\}. \quad (30)$$

3.2. Likelihood function

The likelihood function $\mathcal{L}(\vec{p})$ is calculated from reliability as follows:

$$\mathcal{L}(\vec{p})=\frac{1}{n}\prod_{j=1}^{n}\left(\lambda\left(t_{V_j}^{M}\right)\right)^{\Psi_j}\exp\left\{-\left(\Lambda\left(t_{V_j}^{M}\right)-\Lambda\left(t_{V_{j-1}}^{M}\right)\right)\right\},$$

$$\Psi_j=\begin{cases}1,\text{if event j is a failure / CM.}\\0,\text{if event j is a PM.}\end{cases} \quad (31)$$

The parameter vector $\vec{p}$ is then estimated by the least squares estimation (LSE) method using Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963) implemented in MATLAB.

4. Case study

Based on the maintenance logs and procedures, 3 maintenance actions were identified: thickness adjustment, tightening of a screw and lubrication. At each manoeuvre of a point machine, health indicator is calculated according to Equation 1.

Maintenance effectiveness and other parameters are found for 19 point machines from Italy. The naming convention adopted for the present article is as follows: model ("X" or "Y"), material type used for sliding chairs ("I" or "J"), turnout number and machine sequence letter (1st machine activated in a manoeuvre at a particular turnout is designated as "A", while 2nd machine – as "B").

Observations spanning June 2015–June 2017 were used, with baseline calculated according to the clients' rules to obtain HI values. The data were divided into 3 parts: 'Normal' and 'Reverse', which refer to the operating direction, and 'Both' (the latter combining the former two). Separate parameter estimates were calculated for each. The results, separated by direction, are presented further below.

Maintenance effectiveness estimates with their corresponding 95 % lower and upper confidence limits for the normal direction manoeuvres are presented for 19 point machines in Figure 3. Maintenance effectiveness estimates with their corresponding 95 % lower and upper confidence limits for the reverse direction manoeuvres are presented for 19 point machines in Figure 4. Maintenance effectiveness estimates with their corresponding 95 % lower and upper confidence limits for the combined normal and reverse direction manoeuvres are presented for 19 point machines in Figure 5.

As can be seen from Figures 3-5, the largest variation occurs for the CM effect estimate on virtual age, regardless of the direction of manoeuvres. The variation in the maintenance effect estimates is summarized in Table1.

The maximal age-reducing effect in both directions was most effective maintenance action in both directions is PM effect on virtual age, with the spread of only 13 %, closely followed by the PM effect on virtual health (15 %). For both normal and reverse directions, the smallest spread in estimates is encountered for the PM effect on virtual health (13 % and 8 %, respectively). With $\varphi_{CM}\in[0.01;0.03]$, corrective maintenance appears to have a nearly-AGAN effect on the virtual age for point machine XI10A in all directions of operation. With $\theta_{PM}\in[0.90;0.91]$, preventive maintenance appears to have a nearly-ABAO effect on the virtual HI for point machine XI2A.

For both preventive and corrective maintenance, the majority of the point machines experience imperfect maintenance effects between ABAO and AGAN. Corrective maintenance has an effect closer to that of AGAN on the HI for all point machines in all directions. Corrective maintenance has an effect closer to that of AGAN on the virtual age in 8 out of 19 point machines in both directions, 15 out of 19 point machines in normal direction and 15 out of 19 point machines in reverse direction.

4.1. Estimating the remaining useful life (RUL)

The RUL can be calculated as a pdf:

$$f_{RUL}\left(t_c\mid t_{V_{j-1}}^{M},z_{V_{j-1}}^{+}\right)=\frac{f\left(t+t_{V_j}^{M}-t_{V_{j-1}}^{M}\middle|t_{V_{j-1}}^{M},z_{V_j}^{+}\right)}{R\left(t_{V_j}^{M}-t_{V_{j-1}}^{M}\middle|t_{V_{j-1}}^{M},z_{V_j}^{+}\right)}$$

$$=\lambda\left(t+t_c,z,\varphi_M,\theta_M\mid t_{V_{j-1}}^{M},z_{V_{j-1}}^{+}\right)\frac{R\left(t+t_{V_j}^{M}-t_{V_{j-1}}^{M}\middle|t_{V_{j-1}}^{M},z_{V_{j-1}}^{+}\right)}{R\left(t_{V_j}^{M}-t_{V_{j-1}}^{M}\middle|t_{V_{j-1}}^{M},z_{V_{j-1}}^{+}\right)}, \quad (32)$$

where t is time, T is the lifetime, t_c is the value of RUL random variable $T_c=\left\{t_c:T-t\middle|T\rangle t,t_{V_{j-1}}^{M}\right\}$. The results are shown in Figure 6.

As can be seen from the figure, the predicted RUL is not too far from the actual failure data. The RUL can be predicted without an exact failure threshold based on failure data and condition monitoring (CM) information. The estimated values form a smoother curve than the actual values. This suggests that the estimating procedure is able to smooth the predictions. However, sufficient failure and CM data

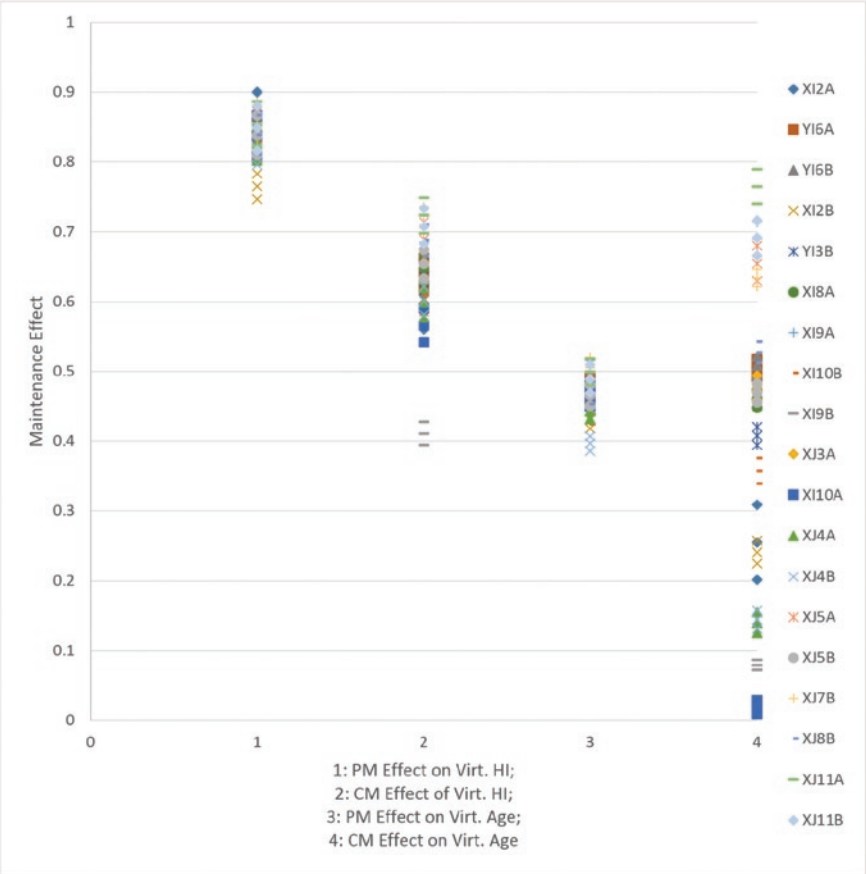


Fig. 3. Maintenance effect on virtual hi and virtual age for point machines in 'normal' direction

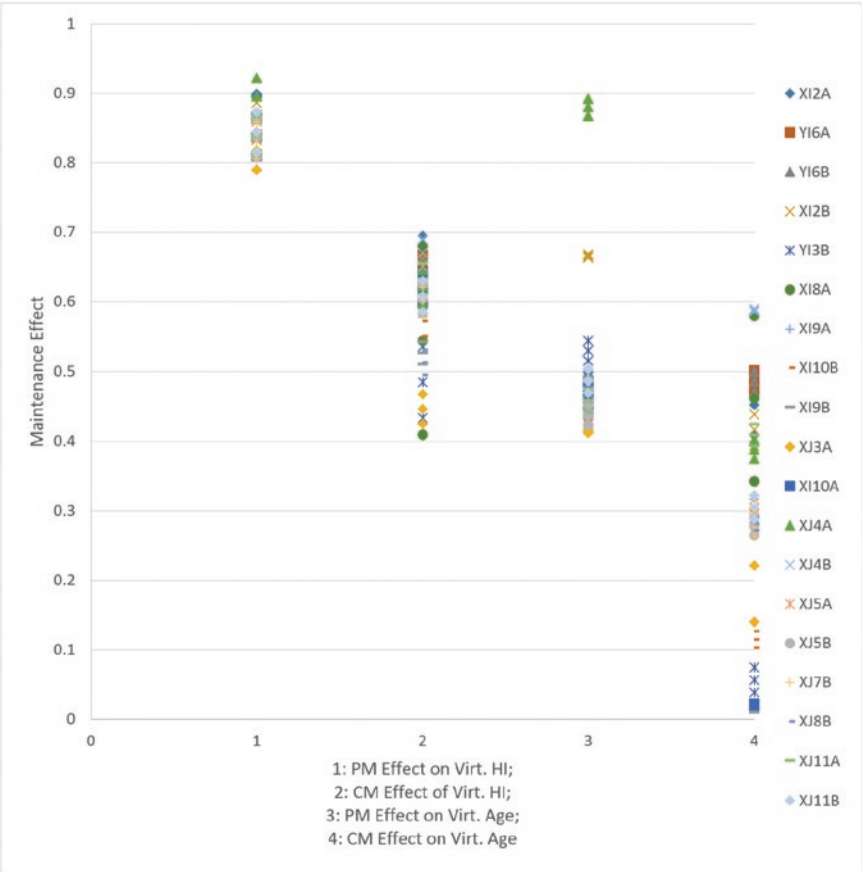


Fig. 4. Maintenance effect on virtual hi and virtual age for point machines in 'reverse' direction

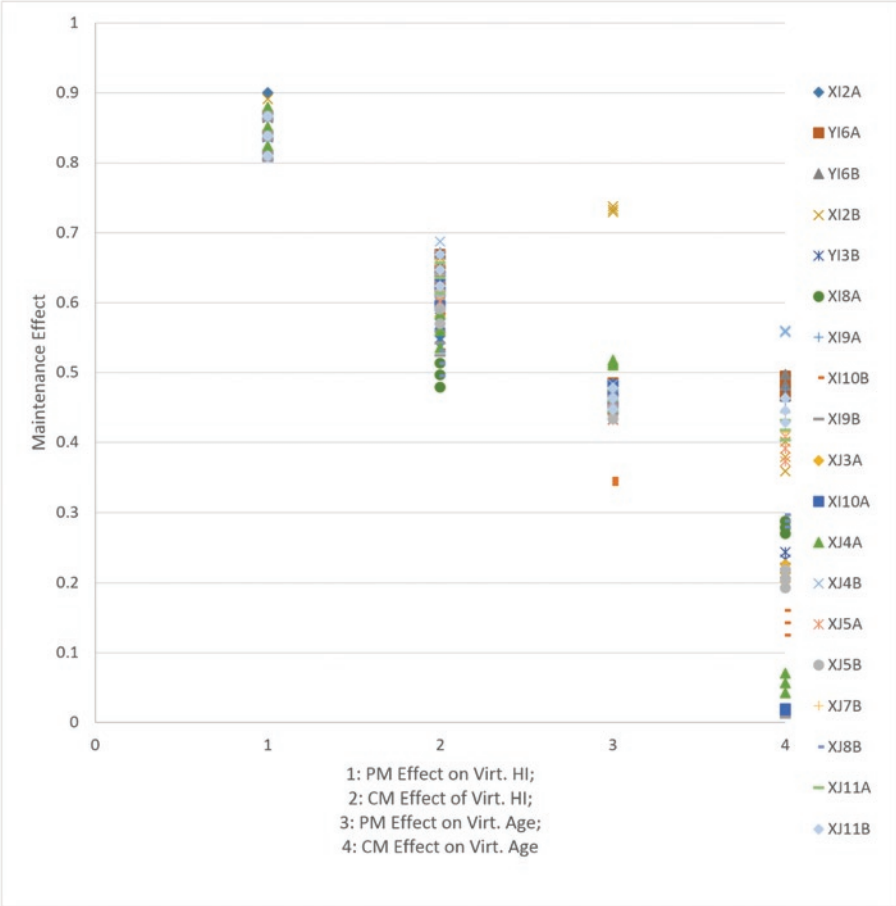


Fig. 5. Maintenance effect on virtual hi and virtual age for point machines in 'both' (i.E. Normal and reverse combined) directions

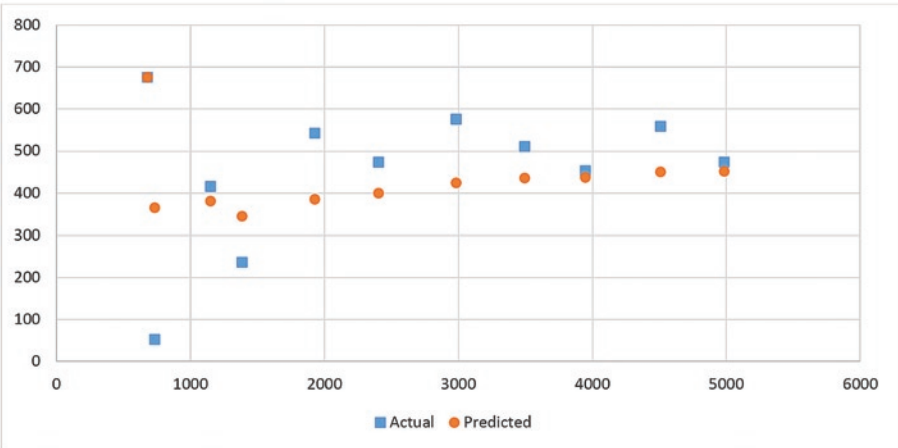


Fig. 6. Actual failures and predicted remaining useful life (rul) estimates

are required, unlike for filtering-based models, where parameters in initial life distribution can be estimated separately.

5. Conclusions

In this paper, a model is proposed for quantifying the effects of different types of maintenance on a device subject to condition monitoring. It is assumed that failures follow a nonhomogeneous Poisson process (NHPP) and covariates follow the Cox proportional hazards model. In particular, the multiplicative effect of maintenance on the age of a device is estimated using the Weibull hazard function, while the multiplicative effect on the health of a device and covariates associated with condition-based monitoring (CBM) is estimated using the Cox hazard function.

The proposed algorithm for estimating the impact and effectiveness of maintenance uses the concept of virtual age and introduces the concept of virtual health. It is shown that virtual health and the effect of maintenance on the health indicator of a device can be described using the concepts of Kalman filter.

An example of practical application of the algorithm is provided to a real case of railway point machines. In this example, preventive or corrective types of maintenance are modelled as different maintenance effect parameters. Using condition monitoring data, the health indicator is calculated as a scaled Mahalanobis distance. The reliability and the likelihood functions are derived and the least squares estimates (LSE) of the covariate coefficient, Weibull shape and scale parameters, as well as the preventive and corrective maintenance effect estimates on time and health indicator are found using the Levenberg-Marquardt algorithm.

The effect of corrective maintenance was closer to that of "as-good-as-new" (AGAN) state across all point machines, with point machine XI10A demonstrating the most dramatic AGAN virtual health improvement. The effect of preventive maintenance on the health indicator was the closest to "as-bad-as-old" (ABAO) across all point machines, with point machine XI2A demonstrating the least improvement in virtual health.

Remaining useful life (RUL) calculations were performed and predicted RUL estimates were obtained. The predicted RUL estimates

Table 1. Estimated Maintenance Effects and Upper and Lower 95 % Confidence Limits [LCL; UCL] on the for 'Both', 'Normal' and 'Reverse' Directions

Direction	PM Effect on Virt. HI, θ_{PM}	CI on PM Effect on Virt. HI, θ_{PM}	CM Effect on Virt. HI, θ_{CM}	CI on CM Effect on Virt. HI, θ_{CM}
Both	0.83	[0.75; 0.90]	0.57	[0.39; 0.75]
Normal	0.86	[0.79; 0.92]	0.55	[0.41; 0.69]
Reverse	0.85	[0.81; 0.89]	0.59	[0.48; 0.69]

Direction	PM Effect on Virt.Age, φ_{PM}	CI on PM Effect on Virt.Age, φ_{PM}	CM Effect on Virt.Age, φ_{CM}	CI on CM Effect on Virt.Age, φ_{CM}
Both	0.46	[0.39; 0.52]	0.4	[0.01; 0.79]
Normal	0.65	[0.42; 0.88]	0.31	[0.02; 0.59]
Reverse	0.54	[0.34; 0.73]	0.29	[0.01; 0.56]

were generally smoother than the actual data, thus displaying filtering qualities.

As a future work, application of fuzzy logic to estimate the health indicator, based on the covariate values appears to be promising. Yet another avenue is to perform clustering analysis using Gaussian mixture model (GMM) and identify the clusters corresponding to normal, failed and/or borderline devices.

Acknowledgements

The authors acknowledge the financial support from the client company for this research. The authors also thank Dr. Pierre Dersin, Dr. Benjamin Lamoureux, Mr. Md Sujauddin Mallick, MEng., Ms. Allegra Alessi, MSc., and Dr. Baptise Labarthe for their help, fruitful discussions and valuable suggestions.

References

1. Ardakani HD, Lucas C, Siegel D, Chang S, Dersin P, Bonnet B, Lee J. PHM for Railway System - a Case Study on the Health Assessment of Point Machines. Proceedings of the Prognostics and Health Management (PHM) IEEE Conference 2012 : 74-79, https://doi.org/10.1109/ICPHM.2012.6299533.
2. Atamuradov V, Medjaher K, Dersin P, Lamoureux B, Zerhouni N. Prognostics and Health Management for Maintenance Practitioners - Review, Implementation and Tools Evaluation. International Journal of Prognostics and Health Management 2017; 8 (Special Issue on Railways & Mass Transportation): 1-31.
3. 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, https://doi.org/10.17531/ein.2018.2.20.
4. Babishin V, Taghipour S. Joint Maintenance and Inspection Optimization of a k-out-of-n System. In: Proceedings of the Annual Reliability and Maintainability Symposium (RAMS) 2016: 1-6, https://doi.org/10.1109/RAMS.2016.7448039.
5. Babishin V, Taghipour S. Joint Optimal Maintenance and Inspection for a k-out-of-n System. International Journal of Advanced Manufacturing Technology 2016;87 (5-8): 1739-1749, https://doi.org/10.1007/s00170-016-8570-z.
6. Babishin V, Taghipour S. Maintenance Effectiveness Estimation with Applications to Railway Industry. In: Proceedings of the Annual Reliability and Maintainability Symposium (RAMS) 2019, https://doi.org/10.1109/RAMS.2019.8769273.
7. 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, https://doi.org/10.1016/j.apm.2016.07.019.
8. Bendell A, Wightman DW, Walker EV. Applying Proportional Hazards Modelling in Reliability. Reliability Engineering and System Safety 1991; 34: 35-53, https://doi.org/10.1016/0951-8320(91)90098-R.
9. Brown M, Proschan F. Imperfect repair. Journal of Applied Probability 1983; 20: 851-859, https://doi.org/10.2307/3213596.
10. Cha JH, Finkelstein M. Optimal Long-Run Imperfect Maintenance With Asymptotic Virtual Age. IEEE Transactions on Reliability 2016; 65(1): 187-196, https://doi.org/10.1109/TR.2015.2451612.
11. Chan JK, Shaw L. Modelling repairable systems with failure rates that depend on age and maintenance. IEEE T Reliab. 1993; 42: 566-570, https://doi.org/10.1109/24.273583.
12. Chen Y, Meng X, Chen S. Reliability Analysis of a Cold Standby System with Imperfect Repair and under Poisson Shocks. Mathematical Problems in Engineering 2014, https://doi.org/10.1155/2014/507846.
13. Conn A, Deleris L, Hosking J, Thorstensen T. A simulation model for improving the maintenance of high cost systems, with application to offshore oil installation. Quality and Reliability Engineering International 2010; 26: 733-748, https://doi.org/10.1002/qre.1136.
14. Corman F, Kraijema S, Godjevac M, Lodewijks G. Optimizing preventive maintenance policy: A data-driven application for a light rail braking system. Proc IMechE Part O: J Risk and Reliability 2017; 231(5): 534-545, https://doi.org/10.1177/1748006X17712662.
15. Cox DR. Regression models and life-tables. New York: Springer; 1992, https://doi.org/10.1007/978-1-4612-4380-9_37.
16. Dagpunar J. Some properties and computational results for a general repair process. Naval Research Logistics 1998; 45: 391-405, https://doi.org/10.1002/(SICI)1520-6750(199806)45:4<391::AID-NAV5>3.0.CO;2-0.
17. Doyen L, Gaudoin O. Classes of imperfect repair models based on reduction of failure intensity or virtual age. Rel Eng and Sys Safety 2004; 84: 45-56, https://doi.org/10.1016/S0951-8320(03)00173-X.
18. Doyen L, Gaudoin O. Imperfect maintenance in a generalized competing risks framework. Journal of Applied Probability 2006; 43: 825-839, https://doi.org/10.1239/jap/1158784949.
19. Fuqing Y, Kumar U. A General Imperfect Repair Model Considering Time-Dependent Repair Effectiveness. IEEE Transactions on Reliability 2012; 61(1): 95-100, https://doi.org/10.1109/TR.2011.2182222.

20. Galar D, Gustafson A, Tormos B, Berges L. Maintenance Decision Making Based on Different Types of Data Fusion. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14(2): 135-144.
21. Grimble MJ. Robust Industrial Control: Optimal Design Approach for Polynomial Systems. New Jersey: Prentice Hall; 1994.
22. Heng A, Tan ACC, Mathew J, Montgomery N, Banjevic D, Jardine AKS. Intelligent condition-based prediction of machinery reliability. Mechanical Systems and Signal Processing 2009; 23: 1600-1614, https://doi.org/10.1016/j.ymssp.2008.12.006.
23. Hu J, Jiang Z, Liao H. Preventive maintenance of a single machine system working under piecewise constant operating condition. Reliability Engineering and System Safety 2017; 168: 105-115, https://doi.org/10.1016/j.ress.2017.05.014.
24. Innotrack. Deliverable 1.4.8 - Overall Cost Reduction. Project Report. Gothenburg, Sweden: Chalmers University of Technology 2009. TIP5-CT-2006-031415.
25. Kallen M. Modelling imperfect maintenance and the reliability of complex systems using superposed renewal processes. Rel Eng and Sys Safety 2011; 96: 636-641, https://doi.org/10.1016/j.ress.2010.12.005.
26. Kalman RE. A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering 1960; 82(35): 35-45, https://doi.org/10.1115/1.3662552.
27. Kijima M. Some results for repairable systems with general repair. Journal of Applied Probability 1989; 26: 89-102, https://doi.org/10.2307/3214319.
28. Kijima M, Nakagawa T. Replacement policies of a shock model with imperfect maintenance. Eur J Oper Res. 1992; 57: 100-110, https://doi.org/10.1016/0377-2217(92)90309-W.
29. Kumar S, Vichare NM, Dolev E, Pecht M. A health indicator method for degradation detection of electronic products. Microelectronics Reliability 2012; 52: 439-445, https://doi.org/10.1016/j.microrel.2011.09.030.
30. Letot C, Dersin P, Pugnaloni M, Dehombreux P, Fleurquin G, Douziech C, La-Cascia P. A Data Driven Degradation-Based Model for the Maintenance of Turnouts: a Case Study. IFAC-PapersOnLine 2015: 958-963, https://doi.org/10.1016/j.ifacol.2015.09.650.
31. Levenberg K. A Method for the Solution of Certain Non-Linear Problems in Least Squares. Applied Mathematics Quarterly 1944; 2: 164-168, https://doi.org/10.1090/qam/10666.
32. Lim TJ, Lie CH. Analysis of system reliability with dependent repair modes. IEEE Transactions on Reliability 2000; 49(2): 153-162, https://doi.org/10.1109/24.877332.
33. Malik M. Reliable preventive maintenance scheduling. AIIE Transactions 1979; 11: 221-228, https://doi.org/10.1080/05695557908974463.
34. Marquardt D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. SIAM Journal on Applied Mathematics 1963; 11(2): 431-441, https://doi.org/10.1137/0111030.
35. Martorell S, Sanchez A, Serradell V. Age-dependent reliability model considering effects of maintenance and working conditions. Reliability Engineering and System Safety 1999; 64(1): 19-31, https://doi.org/10.1016/S0951-8320(98)00050-7.
36. Nasr A, Gasmi S, Sayadi M. Estimation of the parameters for a complex repairable system with preventive and corrective maintenance. In: IEEE Proc, International Conference on Electrical Engineering and Software Applications (ICEESA) 2013: 1-6, https://doi.org/10.1109/ICEESA.2013.6578455.
37. Pham H, Wang H. Imperfect maintenance. Eur J Oper Res. 1996; 94: 425-438, https://doi.org/10.1016/S0377-2217(96)00099-9.
38. Pulcini G. Mechanical Reliability and Maintenance Models. In: H P, editor. Handbook of Reliability Engineering. London: Springer-Verlag 2003: 317-348, https://doi.org/10.1007/1-85233-841-5_18.
39. Said U, Taghipour S. Modeling Failure Process and Quantifying the Effects of Multiple Types of Preventive Maintenance for a Repairable System. Quality and Reliability Engineering International 2016; 33(5): 1149-1161, https://doi.org/10.1002/qre.2088.
40. Syamsundar A, Muralidharan K, Naikan V. General repair models for maintained systems. Sri Lankan Journal of Applied Statistics 2012; 12(1): 117-143, https://doi.org/10.4038/sljastats.v12i0.4971.
41. Syamsundar A, Naikan VNA. Imperfect repair proportional intensity models for maintained systems. IEEE transactions on Reliability 2011; 60(4): 782-787, https://doi.org/10.1109/TR.2011.2161110.
42. Uematsu K, Nishida T. One unit system with a failure rate depending upon the degree of repair. Math. Japonica 1987; 32: 685-691, https://doi.org/10.1016/0026-2714(87)90015-1.
43. Wang Y, Cotofana S. A novel virtual age reliability model for time-to-failure prediction. Integrated Reliability Workshop Final Report (IRW). IEEE; 2010, https://doi.org/10.1109/IIRW.2010.5706498.
44. Wu S, Zuo MJ. Linear and nonlinear preventive maintenance models. IEEE T Reliab. 2010; 59(1): 242-249, https://doi.org/10.1109/TR.2010.2041972.
45. Yu P, Song J, Cassady C. Parameter estimation for a repairable system under imperfect maintenance. In: Proceedings of the Annual Reliability and Maintainability Symposium 2008: 428-433.
46. Zhou X, Xi L, Lee J. Reliability-centred predictive maintenance scheduling for a continuously monitor system subject to degradation. Reliab Eng Syst Safety 2007; 92(4): 530-534, https://doi.org/10.1016/j.ress.2006.01.006.

Vladimir BABISHIN
Sharareh TAGHIPOUR
Department of Mechanical and Industrial Engineering
Ryerson University
350 Victoria Street
Toronto, Ontario, M5B 2K3, Canada

E-Mails: vbabisin@ryerson.ca, sharareh@ryerson.ca

Article citation info:
ČEPIN M. Evaluation of the importance factors of the power plants within the power system reliability evaluation. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 631–637, http://dx.doi.org/10.17531/ein.2019.4.12.

Marko ČEPIN

EVALUATION OF THE IMPORTANCE FACTORS OF THE POWER PLANTS WITHIN THE POWER SYSTEM RELIABILITY EVALUATION

OCENA WSPÓŁCZYNNIKÓW WAŻNOŚCI ELEKTROWNI W RAMACH OCENY NIEZAWODNOŚCI SYSTEMU ELEKTROENERGETYCZNEGO

The objective of the paper is to develop the reliability importance factors which could identify the plants, which more or less contribute to the increased or decreased power system reliability. One group of importance factors could identify the plants, which with their increased availability would notably increase the power system reliability The other group of the importance factors could identify the plants, which with their reduced availability would notably reduce the power system reliability. The importance factors have been developed. An example of regional power system was considered for the case study. The results identify the power plants which are more susceptible to increase of the loss of load expectation and thus to decreasing of the power system reliability, if their availability is reduced. Similarly, the results identify the power plants which are more susceptible to decreasing of the loss of load expectation and thus to increase of the power system reliability, if their availability is increased. The lists of important factors can serve as a standpoint for inclusion of the power system reliability role within the power system to the planning activities. The lists of importance factors can represent the standpoint for the power system operator to reward the improvement of the power plant availability or to penalize the reduced power plant availability.

***Keywords**: power plant, reliability, importance, loss of load expectation, network.*

Celem artykułu jest opracowanie współczynników ważności niezawodności, pozwalających identyfikować elektrownie, które w mniejszym lub większym stopniu przyczyniają się do zwiększenia lub zmniejszenia niezawodności systemu elektroenergetycznego. Opracowano dwie grupy takich współczynników: jedne – służące do identyfikacji elektrowni, które przy zwiększonej gotowości mogą znacznie zwiększać niezawodność systemu elektroenergetycznego, i drugie – do identyfikacji tych elektrowni, których zmniejszona gotowość może znacznie zmniejszać niezawodność sieci elektroenergetycznej. Jako studium przypadku rozważano przykład regionalnego systemu elektroenergetycznego. Wyniki pozwalają na wskazanie elektrowni, które są bardziej podatne na zwiększenie oczekiwanego czasu deficytu mocy (niepokrycia zapotrzebowania), a tym samym mogą bardziej przyczyniać się do zmniejszenia niezawodności systemu elektroenergetycznego, w przypadku spadku ich gotowości. Podobnie, uzyskane wyniki pozwalają ustalić, dla których elektrowni prawdopodobieństwo zmniejszenia oczekiwanego czasu deficytu mocy jest większe, a tym samym, które elektrownie mogą bardziej przyczyniać się do zwiększenia niezawodności systemu elektroenergetycznego, gdy zwiększy się ich gotowość. Listy współczynników ważności mogą służyć jako punkt odniesienia dla działań planistycznych, a także jako podstawa dla operatora systemu elektroenergetycznego do nagradzania poprawy gotowości elektrowni lub penalizacji jej spadku.

***Słowa kluczowe**: elektrownia, niezawodność, ważność, oczekiwany czas niepokrycia zapotrzebowania, sieć.*

1. Introduction

The power system planning is performed considering the parameters related to the investment costs of power plants, considering the parameters related to the operational costs, considering the parameters, which penalise the impact to the environment, but the reliability of the power plants is not normally considered. The idea of the paper is to make a contribution, to rank the power plants in the power system in a way to consider their impact to the reliability of the power system. The classical method of loss of load expectation is selected as the starting point for the development of reliability importance factors, which would rank the power plants according to their reliability performance in the power system [3, 9].

The objective of the paper is to develop the reliability importance factors for the power plants in the power system based on evaluation of the loss of load expectation. One group of the importance factors could identify the plants, which increased availability would notably or significantly increase the power system reliability. Similarly, the other group of the importance factors could identify the plants, which reduced availability would notably or significantly reduce the power system reliability.

State of the art regarding the power systems reliability includes a large number of important methods and studies [2, 5, 8]. Much less has been written regarding the importance factors, although some papers exist in the field [9, 10]. However, they are not based on the loss of load expectation, which is one of the most widely used methods for adequacy evaluation, i. e. determining the reserve power for the static power system reliability evaluation [3, 5-7, 9, 15, 18]. Section 2 gives the overview of existing methods and newly developed importance factors. Section 3 gives the results of the case study showing the values of newly developed importance factors on selected cases. Section 4 gives conclusions.

2. Methods

2.1. The LOLE method

The focus of the paper is placed to the static methods aiming at power system adequacy [2, 3, 8, 9]. The loss of load expectation

(LOLE) is a standardised method for determination of the power system reserve in order that the power system reliability does not fall below the required reliability level.

LOLE represent the number of hours in a year (other units are possible), in which the power consumption could not be covered by the available power plants. Many states are considered. Each state expresses a configuration of the power system related to information that some power plants are available and some are unavailable at certain moment. The states are related to the load diagram. The load diagram is a diagram, which represent the hourly distributed load of all the consumers, which needs to be supplied [2, 3, 8, 9].

Figure 1 shows the parameters for evaluation of the LOLE on an example load diagram. The example load diagram is not ordered by the actual daily time, but the time points are followed by the decreasing capacity, as it is needed for evaluation of LOLE.

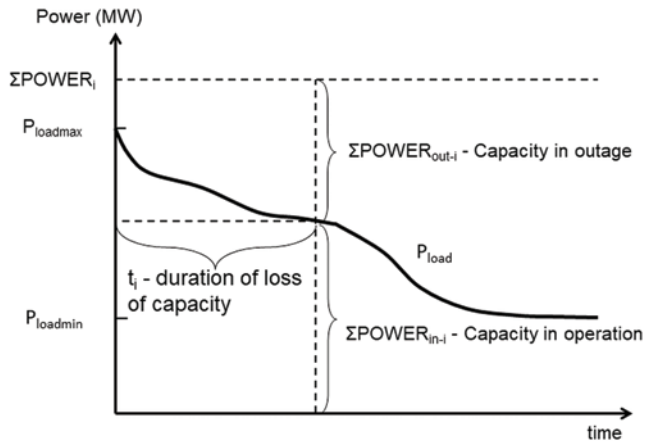


Fig. 1. Graphical representation for determining the loss of load expectation

The equation for calculation of LOLE:

$$LOLE = \sum_{i=1}^{ii} p_i \cdot t_i \tag{1}$$

$$t_i = time \qquad where \qquad P_{load} > \sum POWER_{in-i} \tag{2}$$

i – index of considered state

p_i – probability of state i

t_i – duration of loss of capacity of state i, the time interval, in which the capacity of power plants in operation does not reach the power of load

ii – the number of states

$ii = 2^n$

n – number of power plants in the system

P_{load} – power system load with its minimal value $P_{loadmin}$ and its maximal value - $P_{loadmax}$

$POWER_{in-i}$ – the sum of powers of the power plants, which are assumed available in state i

$POWER_{out-i}$ – the sum of powers of the power plants, which are assumed unavailable in state i

Probability of state i is evaluated as a product of availabilities and unavailabilities of plants assumed available or unavailable in this state. Thus, the states differ among each other by considering different sets of available and unavailable plants. In the case of one plant: it can be available or unavailable. In the case of two plants, we have four states: both available, both unavailable, first available and second unavailable, first unavailable and second available.

$$p(i) = \prod_{r=1}^{n1} a(r) \cdot \prod_{s=1}^{n2} \left(1 - a(s)\right) \tag{3}$$

n1 – number of plants available for certain state

n2 – number of plants unavailable for certain state

n=n1+n2

a(r) – availability of plant r (unavailability is expressed as u(r)=1-a(r) is its complement, or availability is expressed with forced outage rate (FOR) as a(r)=1-FOR(r), because forced outage rate can represent unavailability, u(r)= FOR(r))

a(s) – availability of plant s

The number of states and thus the time of evaluation increases significantly with the increase of number of power plants considered. Therefore, recursive algorithms have been developed, which does not go state by state when evaluating the LOLE, but they build the evaluation by adding the plants one by one to the evaluation [15, 18, 29, 31, 32].

Table 1 shows a small example of a power system including the data and the results, which give LOLE. The load diagram has the maximum power of 40 MW and the minimal power of 10 MW and in between it is linear. Figure 2 shows the load diagram.

Explanation of the table says that the first line shows state 1 and all three units in operation. Their total unavailable power is 0 MW and the total available power is 80 MW. The product of plant availabilities gives the probability of the state 0.8208. Their power is all the time larger than the load, so the time duration of loss of capacity is 0. The subsequent lines follow the described method. Figure 2 shows that 8 hours are such, where capacity service of 30 MW is smaller than the load.

Table 1. Example of capacity table for three power plants.

State k	Unit A 40 MW a(A)=0.9	Unit B 30 MW a(B)=0.95	Unit C 10 MW a(C)=0.96	Capacity lost (MW)	Capacity in service (MW)	Probability of each capacity state, p(k)	t_{loss}(k) (h/d)	p(k)* t_{loss}(k) (h/d)
1	1	1	1	0	80	0.9·0.95·0.96=0.8208	0	0
2	1	1	0	10	70	0.90·0.95·0.04=0.0342	0	0
3	1	0	1	30	50	0.90·0.05·0.96=0.0432	0	0
4	0	1	1	40	40	0.10·0.95·0.96=0.0912	0	0
5	1	0	0	40	40	0.90·0.05·0.04=0.0018	0	0
6	0	1	0	50	30	0.10·0.95·0.04=0.0038	8	0.0304
7	0	0	1	70	10	0.10·0.05·0.96=0.0048	24	0.1152
8	0	0	0	80	0	0.01·0.05·0.04=0.0002	24	0.0048
						LOLE (hours/day) =		0.1504

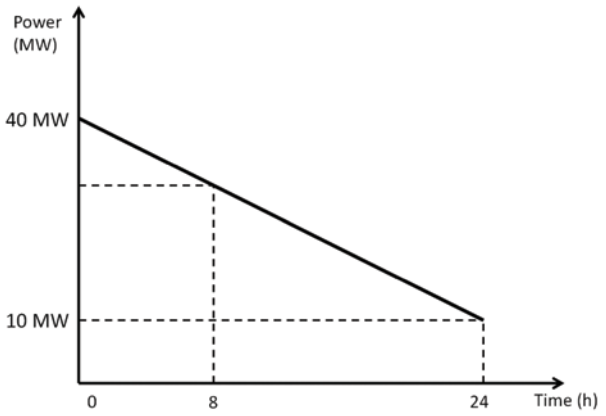


Fig. 2.Load diagram for a small example system

2.2. The improved LOLE method

The improved loss of load expectation method has been developed keeping in mind that the power plants, which rely on environmental conditions cannot have the nominal power available all the time [1, 4, 11, 13-14, 16]. Their power changes with regard to the related parameters such as river flow in the case of hydro power plants, wind speed in the case of wind power plants and sun irradiance in the case of the solar power plants [5, 6, 8]. Please note, that the listed factors are only representative for each plant. In reality, much more factors are involved in mathematical models, where the power of the specific plant is expressed based on influencing factors such as pressure, temperature and humidity for example. Figure 3 shows graphical representation of parameters for determining the improved LOLE. The figure shows that the overall considered time interval (usually a day or a year) is divided to 6 time windows for the figure being indicative. In reality, the considered time windows are cut to 1 hour windows even to time windows of shorter time durations.

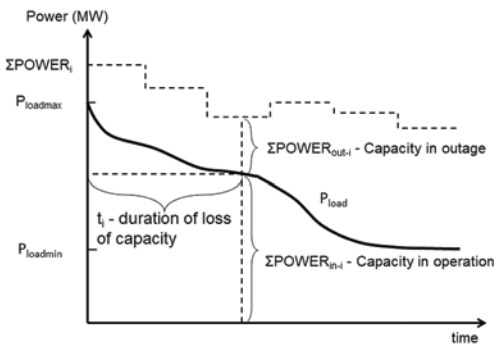


Fig. 3. Graphical representation of parameters for determining improved loss of load expectation

Evaluation of the loss of load expectation is calculated for each time window, its mean value over all time windows can give the average:

$$LOLE = \frac{i}{z}\sum_{j=1}^{z} LOLE_j \tag{4}$$

$LOLE_j$ – loss of load expectation in time window j (hours per day)
LOLE – loss of load expectation (hours per day)
z = number of time windows considered
j = index of a time window

2.3. Importance factors

The importance factors: RIF (risk increase factor) and RDF (risk decrease factor) have been originally developed for expressing the importance of different safety equipment in the safety and reliability analyses [9, 33]. Some other importance factors and their variants followed [10, 12, 20]. RIF was initially named as risk achievement worth (RAW). RDF was initially named as risk reduction worth (RRW).

The equation for calculation of RAW is the following:

$$RAW_a = \frac{P_s\left(P_a = 1\right)}{P_s} \tag{5}$$

RAW_a – risk achievement worth for failure of component A modeled in event a,
$P_s(P_a=1)$ – system failure probability when failure probability of component A modeled in event a is set to 1,
P_s – system failure probability.

The equation for calculation of RRW is the following:

$$RRW_a = \frac{P_s}{P_s\left(P_a = 0\right)} \tag{6}$$

RRW_a – risk reduction worth for failure of component A modeled in event a,
$P_s(P_a=0)$ – system failure probability when failure probability of component A modeled in event a is set to 0,
P_s – system failure probability.

2.4. New importance factors

The importance factors are factors, which based on some determined sensitivity analyses determine the impact of a certain parameter to the calculated result. They help ranking the components of the system based on the selected parameters.

The following importance factors have been developed for ranking of the power generating plants in the power system in sense to judge their role in contributing to the static system reliability and in sense to judge their role in contributing to the determining the power reserve needed for assuring power system reliability.

LOLE absolute increase factor for power plant i shows how much LOLE is increased in the case if the power plant i would become totally unavailable (its availability becomes 0, a(i)=0):

$$\mathrm{LOLE}_{\mathrm{abs-increase-i}} = \frac{LOLE_{a(i)=0}}{LOLE} \tag{7}$$

LOLE absolute decrease factor for power plant i shows how much LOLE is decreased in the case if the power plant i would become totally available (its availability becomes 1, a(i)=1):

$$\mathrm{LOLE}_{\mathrm{abs-decrease-i}} = \frac{LOLE}{LOLE_{a(i)=1}} \tag{8}$$

LOLE – loss of load expectation for the power system considering the number of power plants in the system and considering the load diagram
$LOLE_{abs\text{-}increase\text{-}i}$ – LOLE absolute increase factor for the power plant i
$LOLE_{abs\text{-}decrease\text{-}i}$ – LOLE absolute decrease factor for the power plant i
$LOLE_{a(i)=0}$ – LOLE if availability of the power plant i equals to 0
$LOLE_{a(i)=1}$ – LOLE if availability of the power plant i equals to 1

When LOLE is increased in the power system, this means more hours per year without power supply, and it is related with decrease of the power system reliability. Plants with high $LOLE_{abs\text{-}increase}$ are important in the power system in order to maintain them and monitor their availability in order not to become unavailable. Reduction of their availability affects the system reliability in a large extent.

Plants with high $LOLE_{abs\text{-}decrease}$ are important in the power system for its reliability improvement. Their improved availability in a large extent causes LOLE decrease and thus the power reliability increase.

Those two importance factors: $LOLE_{abs\text{-}increase\text{-}i}$ and $LOLE_{abs\text{-}decrease}$ can be considered as static, similarly as it is LOLE in section 2.1.

When considering different $LOLE_j$ in different time windows (see eq. 4 in section 2.2), one has to develop also importance factors for each time window: $LOLE_{abs\text{-}increase\text{-}i\text{-}j}$ and $LOLE_{abs\text{-}decrease\text{-}i\text{-}j}$.

Each of the the time windows with specific power system conditions need to be considered separately for better representation of the system behavior. In addition the average can be calculated as for LOLE in eq. 4.

The different LOLE absolute increase factors for the time windows can give also their mean value and show the change of this factor at different time windows of the power system. Similarly, the different LOLE absolute decrease factors for the time windows can give their mean value and show the change of this factor at different time windows of the power system:

$$\mathrm{LOLE}_{\mathrm{abs-increase-i-mean}} = \frac{1}{jj}\sum_{j=1}^{jj}\mathrm{LOLE}_{\mathrm{abs-increase-i-j}} = \frac{1}{jj}\sum_{j=1}^{jj}\frac{LOLE_{j\_a(i)=0}}{LOLE_j} \quad (9)$$

$$\mathrm{LOLE}_{\mathrm{abs-decrease-i-mean}} = \frac{1}{jj}\sum_{j=1}^{jj}\mathrm{LOLE}_{\mathrm{abs-decrease-i-j}} = \frac{1}{jj}\sum_{j=1}^{jj}\frac{LOLE_j}{LOLE_{j\_a(i)=1}} \quad (10)$$

$LOLE_{abs\text{-}increase\text{-}i\text{-}mean}$ – mean LOLE absolute increase factor for the power plant i

$LOLE_{abs\text{-}decrease\text{-}i\text{-}mean}$ – mean LOLE absolute decrease factor for the power plant i

$LOLE_{abs\text{-}increase\text{-}i\text{-}j}$ – LOLE absolute increase factor for the power plant i in time window j

$LOLE_{abs\text{-}decrease\text{-}i\text{-}j}$ – LOLE absolute decrease factor for the power plant i in time window j

$LOLE_{j\_a(i)=0}$ – LOLE in the time interval j if availability of the power plant i equals to 0

$LOLE_{j\_a(i)=1}$ – LOLE in the time interval j if availability of the power plant i equals to 1

$LOLE_j$ – loss of load expectation for the power system in the time interval j

3. Analyses and Results

3.1. Models

Several power system models were used for application of the developed models. The verification of the algorithm for LOLE was performed for reliability test system.

An example of regional power system was considered for the case study. The selected regional system includes nuclear power plant, thermal power plants, hydro power plants and wind power plant. Table 2 shows the characteristics of the regional power system, which are needed for the evaluation of LOLE and the importance factors.

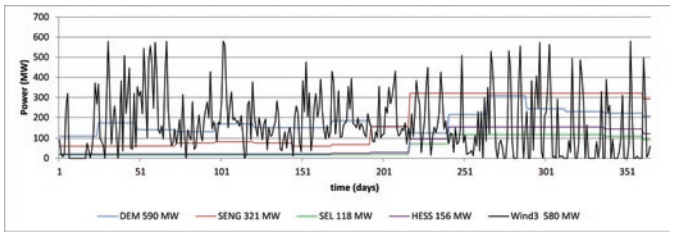


Fig. 4. The timely curves of the power of the power plants in the power system, which are subjected to the power system changes versus time

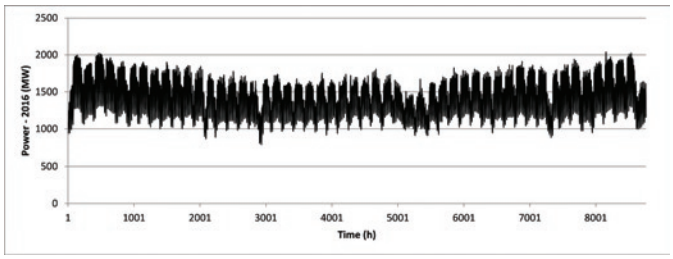


Fig. 5 The power consumption in the power system in the year 2016

Table 2. Characteristics of the regional power system

Plant Identification (includes nominal power)	Plant Type	Forced outage rate (FOR); unavailability	1-FOR; availability
NEK 696 MW	Nuclear	0.01	0.99
TEŠ6 544 MW	Thermal (coal)	0.08	0.92
TEŠ5 345 MW	Thermal (coal)	0.09	0.91
TEŠ4 275 MW	Thermal (coal)	0.09	0.91
TEŠplin 84 MW	Thermal (gas)	0.06	0.94
TETOL 124 MW	Thermal (coal)	0.06	0.94
TEB 350 MW	Thermal (gas)	0.04	0.96
DEM 590 MW	Hydro	0.01	0.99
SENG 321 MW	Hydro	0.01	0.99
SEL 118 MW	Hydro	0.01	0.99
HESS 156 MW	Hydro	0.01	0.99
Biomass 442 MW	Biomass	0.01	0.99
Wind 580 MW	Wind	0.01	0.99

Figure 4 shows the power plants, which are subjected to the power changes due to the weather parameters such as hydro power, which depends on the river flow and wind power, which depends on the wind speed.

Figure 5 shows the power consumption in the power system in the year 2016 including the losses at the transmission and distribution.

3.2. Analyses and results

LOLE was calculated for the example regional power system. Actually, LOLE was calculated in every time window (every day), considering the actual power plant power at each time window related to the actual weather parameters (river flow, wind speed) instead of considering the nominal power through all the evaluation.

Table 3 shows the calculated average importance factors for the regional power system. Figure 6 shows the LOLE absolute increase factors for the plants in the example power system. Figure 7 shows the LOLE absolute decrease factors for the plants in the example power system.

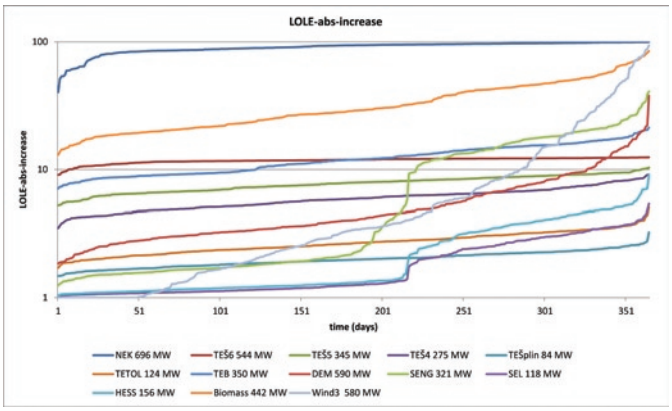


Fig. 6. The LOLE absolute increase factors for the plants in the example power system

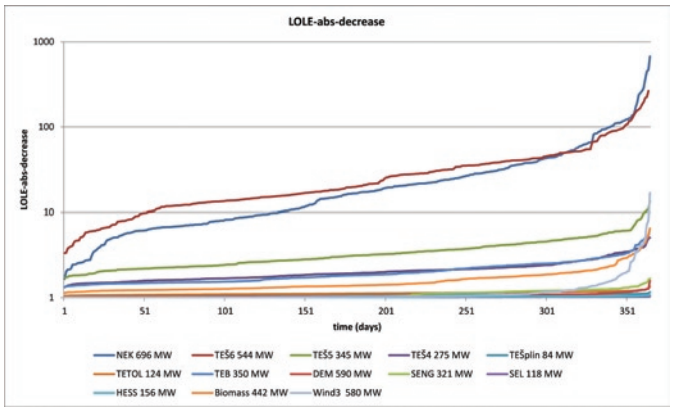


Fig. 7. The LOLE absolute decrease factors for the plants in the example power system

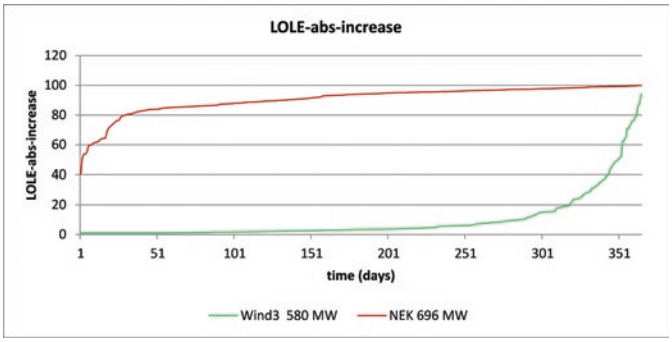


Fig. 8. The LOLE absolute increase factors for the plants in the example power system – nuclear versus wind

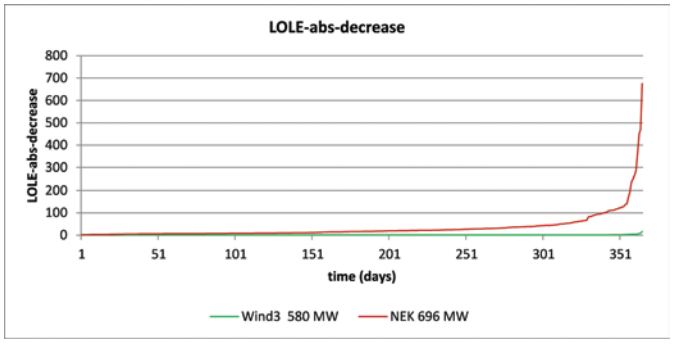


Fig. 9. The LOLE absolute decrease factors for the plants in the example power system – nuclear versus wind

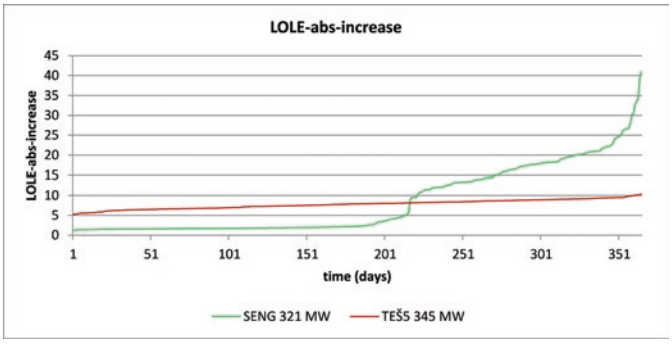


Fig. 10. The LOLE absolute increase factors for the plants in the example power system – thermal versus hydro

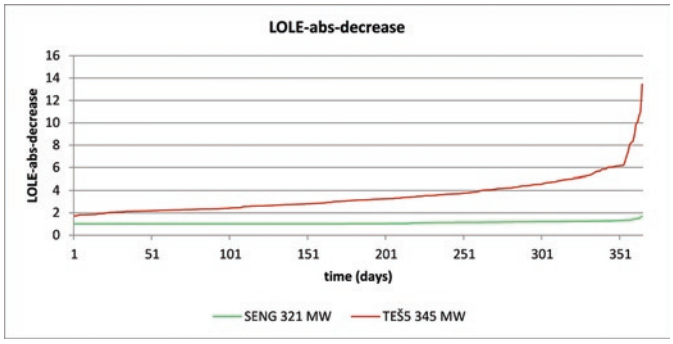


Fig. 11. The LOLE absolute decrease factors for the plants in the example power system – thermal versus hydro

Table 3. Average importance factors for the regional power system

Plant Identification (includes nominal power)	Forced outage rate (FOR)	$LOLE_{abs\text{-}increase\text{-}mean}$	$LOLE_{abs\text{-}decrease\text{-}mean}$
NEK 696 MW	0.01	90.76	33.66
TEŠ6 544 MW	0.08	11.80	32.15
TEŠ5 345 MW	0.09	7.79	3.50
TEŠ4 275 MW	0.09	5.95	2.08
TEŠgas 84 MW	0.06	2.00	1.06
TETOL 124 MW	0.06	2.72	1.12
TEB 350 MW	0.04	12.29	2.04
DEM 590 MW	0.01	5.64	1.05
SENG 321 MW	0.01	8.41	1.09
SEL 118 MW	0.01	1.86	1.00
HESS 156 MW	0.01	2.34	1.01
Biomass1 442 MW	0.01	33.49	1.61
Wind1 580 MW	0.01	9.66	1.23

Figure 8 shows the comparison the nuclear power and wind power related to the LOLE absolute increase factors. Figure 9 shows the comparison the nuclear power and wind power related to the LOLE absolute decrease factors.

Figure 10 shows the comparison the hydro power and thermal power related to the LOLE absolute increase factors. Figure 11 shows the comparison the hydro power and thermal power related to the LOLE absolute increase factors.

The results show that the power plants which can in the largest extent contribute to the decrease of the power system reliability (if their availability is reduced) are the nuclear (NEK), the thermal (TEŠ6 and TEB) and the biomass (Biomass1) power plants. Those plants availability need to be kept at the current level, otherwise the reliability of power system can be largely reduced (i.e. LOLE largely increased).

The results show that the power plants which can in the largest extent contribute to the increase

of the power system reliability (if their availability is improved) are the nuclear (NEK) and the thermal (TEŠ6) power plants. All others can contribute to better power system reliability in much smaller extent. Those two plants need to be mostly credited for the good power system reliability. The system reliability measured by LOLE is 0.43 hours per year (average through all time windows).

The results can represent the standpoint for planning of future power systems. Normally, the optimisation of the levelised costs of electricity include the investment costs, the operational costs the impact to the environment, but no contribution to the reliability of power system. Application of this method could mean the standpoint for inclusion of the power system reliability role within the power system to the planning activities in order that some renewable plants are not favorited too much.

4. Conclusion

The objective related to development of the reliability importance factors for power plants in the system has been reached. New reliability importance factors have been developed. Their application on the real regional power system example was demonstrated. The results identify the most important plants in terms of mean LOLE absolute increase factor and thus decreasing the power system reliability. Similarly, the results identify the most important plants in terms of mean LOLE absolute decrease factor and thus increasing the power system reliability.

The lists of important factors can serve as a standpoint for inclusion of the power system reliability role within the power system to the planning activities.

The lists of importance factors can represent the standpoint for the power system operator to reward the improvement of the power plant availability or to penalize the reduced power plant availability.

Further work can be directed to develop the quantitative specifications of the importance factors for the power system planning and to develop the quantitative criteria for rewarding the availability improvement at each plant and the criteria for penalising the availability reduction at each plant.

Acknowledgement

The Slovenian Research Agency supported partially this research within the research program P2-0356.

References

1. Abdullah M, Muttaqi K, Agalgaonkar A P, Sutanto D. A noniterative method to estimate load carrying capability of generating units in a renewable energy rich power grid. IEEE Transactions on Sustainable Energy 2014; 5(3): 854-865, https://doi.org/10.1109/TSTE.2014.2307855.
2. Anders G J. Probability concepts in electric power systems. John Wiley and Sons, 1989.
3. Billinton R, Allan R. Reliability evaluation of power systems. Plenum Press, 1996, https://doi.org/10.1007/978-1-4899-1860-4.
4. Brancucci Martínez-Anidoa C, Bolado R, De Vries L, Fulli G, Vandenbergh M, Masera M. European power grid reliability indicators, what do they really tell? Electric Power Systems Research 2012; 90: 79-84, https://doi.org/10.1016/j.epsr.2012.04.007.
5. Bricman Rejc Ž, Čepin M. An improved method for power system generation reliability assessment (in Slovenian). Electrotechnical Review 2013; 80(1/2): 57-63.
6. Bricman Rejc Ž, Čepin M. Estimating the additional operating reserve in power systems with installed renewable energy sources. International Journal of Electrical Power & Energy Systems 2014; 62: 654-664, https://doi.org/10.1016/j.ijepes.2014.05.019.
7. Calabrese G. Generating reserve capacity determined by the probability method. American Institute of Electrical Engineers Transactions 1947; 66: 1439-50, https://doi.org/10.1109/T-AIEE.1947.5059596.
8. Čepin M. Reliability of power system considering replacement of conventional power plants with renewables. Safety and reliability - safe societies in a changing world: Proceedings of the 28th European Safety and Reliability Conference (ESREL 2018), 2018. Boca Raton, CRC Press, Taylor & Francis Group, 63-70, https://doi.org/10.1201/9781351174664-8.
9. Čepin M. Assessment of power system reliability. Springer, 2011, https://doi.org/10.1007/978-0-85729-688-7.
10. Čepin M, Volkanovski A. New importance factors in electric power systems (in Slovenian), Electrotechnical Review 2009; 76(4): 177-181.
11. Dehghan S, Kiani B, Kazemi A, Parizad A. Optimal Sizing of a Hybrid Wind/PV Plant Considering Reliability Indices. World Academy of Science, Engineering and Technology International Journal of Electrical and Computer Engineering 2009; 3(8): 1546-1554.
12. Duflot N, Bérenguer C, Dieulle L, Vasseur D., A min cut-set-wise truncation procedure for importance measures computation in probabilistic safety assessment, Reliability Engineering & System Safety 2009, 94 (11): 1827-1837, https://doi.org/10.1016/j.ress.2009.05.015.
13. Elmakias D. New computational methods in power system reliability. Springer Verlag Berlin Heidelberg, 2008.
14. Gami D. Effective load carrying capacity of solar PV plants: a case study across USA. The Ohio State University, Master Thesis, 2016.
15. Garver L L. Effective load carrying capability of generating units, Transactions on Power Apparatus and Systems 1996; 85(8): 910-919, https://doi.org/10.1109/TPAS.1966.291652.
16. Gjorgiev B, Kančev D, Čepin M. A new model for optimal generation scheduling of power system considering generation units availability. International Journal of Electrical Power and Energy Systems 2013; 47(1): 129-139, https://doi.org/10.1016/j.ijepes.2012.11.001.
17. IEEE Std 1366. Guide for electric power distribution reliability indices. IEEE, 2003.
18. Kirn B, Čepin M, Topič M. Effective load carrying capability of solar photovoltaic power plants - case study for Slovenia. Safety and reliability: theory and applications: Proceedings of the 27th European Safety and Reliability Conference. Taylor and Francis, 2017: 3231-3239, https://doi.org/10.1201/9781315210469-408.
19. Kolenc M, Papič I, Blažič B. Coordinated reactive power control to ensure fairness in active distribution grids. International Conference-Workshop Compatibility and Power Electronics 2013: 109-114, https://doi.org/10.1109/CPE.2013.6601138.
20. Langeron Y, Barros A, Grall A, Bérenguer C., Dependability assessment of network-based safety-related system, Journal of Loss Prevention in the Process Industries 2011, 24 (5): 622-631, https://doi.org/10.1016/j.jlp.2011.05.008.
21. Mancarella P, Puschel S, Zhang L, Wang H, Brear M, Jones T, Jeppesen M, Batterham R, Evans R, Mareels I. Power System Security Assessment of the future National Electricity Market. University of Melbourne, 2017.
22. Melhorn A C. Unit commitment methods to accommodate high levels of wind generation. University of Tennessee, Master Thesis, 2011.
23. Mihalič R, Povh D, Pihler J. Stability and dynamic phenomena in power systems (in Slovenian: Stabilnost in dinamični pojavi v elektroenergetskih sistemih). CIGRE CIRED, 2013.

24. Milligan M, Porter K. Determining the capacity value of wind: an updated survey of methods and implementation. National Renewable Energy Laboratory 2008: 1-26.
25. Omahen G, Blažič B, Kosmač J, Souvent A, Papič I. Impact of the SmartGrids concept on future distribution system investments in Slovenia. CIRED 2012 Workshop: Integration of Renewables into the Distribution Grid 2012: 1-4, https://doi.org/10.1049/cp.2012.0831.
26. Pantoš M. Stochastic generation-expansion planning and diversification of energy transmission paths. Electric Power Systems Research 2013; (98): 1-10, https://doi.org/10.1016/j.epsr.2012.12.017.
27. Paska J. Chosen aspects of electric power system reliability optimization. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(2): 202-208.
28. Perkin S, Svendsen A B, Tollefsen T, Honve I, Baldursdottir I, Stefansson H, Kristjansson R, Jensson P. Modelling weather dependence in online reliability assessment of power systems. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2017; 231(4): 364-372, https://doi.org/10.1177/1748006X17694951.
29. Phoon H Y. Generation System Reliability Evaluations with Intermittent Renewables. Master Thesis, University of Strathclyde, 2006.
30. Rosinski A, Dabrowski T. Modelling reliability of uninterruptible power supply units, Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(4): 409-413.
31. Wang X F, McDonald J, Xifan W, Wang X F. Modern power system planning. McGraw-Hill, 1994.
32. Wangdee W, Li W, Billinton R. Pertinent factors influencing an effective load carrying capability and its application to intermittent generation. International Journal of Systems Assurance Engineering and Management 2010; (1)2: 146-156, https://doi.org/10.1007/s13198-010-0025-6.
33. Vesely W E, Belhadj M, Rezos J T. PRA importance measures for maintenance prioritization applications, Reliability Engineering & System Safety 1994, 43(3): 307-318, https://doi.org/10.1016/0951-8320(94)90035-3.

Marko ČEPIN
Faculty of Electrical Engineering
University of Ljubljana
Tržaška cesta 25, 1000 Ljubljana, Slovenia

E-mail: marko.cepin@fe.uni-lj.si

Article citation info:
ZAHARIA SM. The methodology of fatigue lifetime prediction and validation based on accelerated reliability testing of the rotor pitch links. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 638–644, http://dx.doi.org/10.17531/ein.2019.4.13.

Sebastian Marian ZAHARIA

THE METHODOLOGY OF FATIGUE LIFETIME PREDICTION AND VALIDATION BASED ON ACCELERATED RELIABILITY TESTING OF THE ROTOR PITCH LINKS

METODOLOGIA PROGNOZOWANIA TRWAŁOŚCI ZMĘCZENIOWEJ ORAZ JEJ WALIDACJA W OPARCIU O PRZYSPIESZONE BADANIA NIEZAWODNOŚCI DŹWIGNI SKOKU WIRNIKA NOŚNEGO

Because the industrial products have lifetimes, without failing, of up to millions of cycles, it is mandatory that the aerospace field puts into practice the accelerated testing techniques. The lifetime prediction methodology for industrial products presented in this paper was put into practice by performing accelerated reliability testing on an aerospace product (the pitch link of a helicopter). The results showed a significant reduction of the testing time and costs. One important aspect highlighted in this paper is the equivalence between accelerated reliability testing and the traditional reliability testing, by using the two fundamental principles of the accelerated experiments: first, the stresses applied must not alter the physical mechanism through which the defects are produced and second, the conservation of the distribution laws of the failure times. In this way, by equivalence of the accelerated experiments, the methodology contained in this paper was validated.

***Keywords**: reliability, validation, fatigue life, accelerated reliability testing, pitch links.*

Ponieważ okresy bezawaryjnego użytkowania produktów przemysłowych stosowanych w w branży lotniczej mogą wynosić nawet kilka milionów cykli, badanie niezawodności tych wyrobów wymaga zastosowania technik badania przyspieszonego. Metodologię prognozowania czasu pracy produktów przemysłowych przedstawioną w niniejszym artykule wykorzystano w badaniach przyspieszonych niezawodności dźwigni skoku wirnika nośnego helikoptera. Wyniki wykazały, że proponowana metoda pozwala na znaczną redukcję czasu i kosztów badania. Ważnym aspektem, podkreślonym w niniejszej pracy, jest równoważność przyspieszonych i tradycyjnych badań niezawodności, którą można uzyskać respektując dwie podstawowe zasady eksperymentów przyspieszonych: po pierwsze, zastosowane naprężenia nie mogą zmieniać fizycznego mechanizmu, który prowadzi do powstania wady, a po drugie, należy przestrzegać praw dotyczących rozkładu czasów uszkodzeń. Przeprowadzone badania potwierdzają poprawność proponowanej metody.

***Słowa kluczowe**: niezawodność, walidacja, trwałość zmęczeniowa, przyspieszone testy niezawodności, dźwignia skoku.*

1. Introduction

The growing global competition determined the producers to develop products having multiple characteristics with high reliability, at a reduced cost and in the shortest time possible. The challenges posed by these objectives pushed forward the manufacturers to develop and use efficient reliability methods that include accelerated reliability testing. Accelerated experiments are an economic way of getting faster the information regarding the behavior of the products.

The acceleration of the conditions, meaning the "time testing compression" can be studied relative to the number of cycles until failure. To reduce the testing time the stress is applied over the normal limits, keeping the mechanism of failure [9].

The accelerated reliability testing imposes limits like [6,7]:

- the nature of the defects for the accelerated levels has to be the same;
- the test specimens subjected to accelerated testing have to be similar to those used for normal stress;
- the adjustment of the testing model has to be in accordance with the tested product's working parameters;
- every sampling tested at a certain stress needs to be statistically homogenous;
- the results of the accelerated testing must not be extrapolated outside the boundaries of the acceleration model.
- the acceleration model between stress and life time has to conform structurally and functionally to the tested product.
- the accelerated levels must not modify the way the product fails in normal conditions (the distribution of the operating time is not modified, meaning that the shape of the probability density is not changed).

The accelerated reliability tests are developed in a great variety. Each company has the freedom to choose the applied loads for its products because these are considered internal tests, the client receiving only the equivalent results (reliability indicators) determined by extrapolating them from the accelerated level to the normal use one [11]. The research regarding the failures of the aerospace structures highlight the necessity to complement and implement modern computational methods for testing them, both in static and dynamic level.

The fatigue tests have decisive influence over the reliability of aerospace structures so that the statistical characteristic of the fatigue calculation for this kind of stress has to be taken into consideration, including the statistical nature of the stress itself. If an aerospace structure requires for example 10^6-10^7 cycles to produce a fatigue failure

in normal testing conditions, by using of accelerated testing the same result can be obtained after 10^4-10^5 cycles [14,16, 18].

In the case study (the helicopter pitch link) presented in this paper there a cyclical mechanical stress was applied. For cyclical stresses the mechanical systems and components are most often used and the most often met failing phenomenon is the fatigue. The fatigue testing for different components (helicopter blade [20], supple platinum [21], pitch links [17,23], wing spar [3,13] and landing gear [1,12]) can have millions of cycles until failure. For this reason the use of accelerated testing is a method through which the time testing for aerospace components is shortened and thus the testing system is made more efficient [5, 8, 10, 22]. The scope of the present study is the investigation of fatigue life prediction of pitch links components from the IAR 330 helicopter structure subjected to accelerated reliability testing.

2. Experimental details

In fig. 1 is presented: the anti-torque rotor hub contains a body (2) and five assemblies spindle-sleeve (3) that allows: blade feathering; blade pitch change through the swash plate (1) and the pitch link (4) to the swash plate – pitch lever (5).

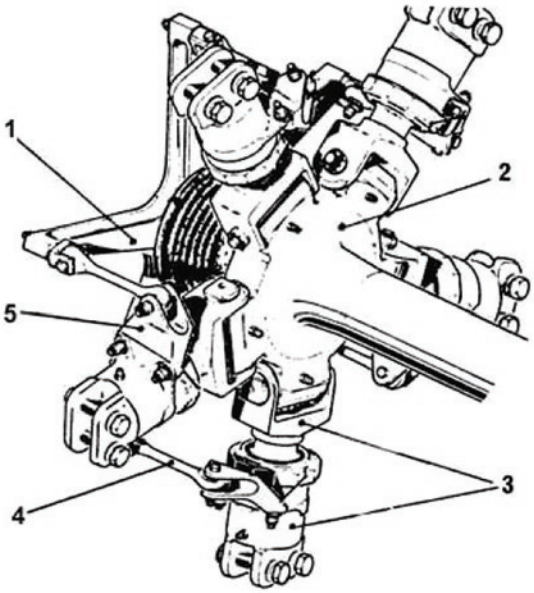


Fig. 1. Rotor hub assembly [4]

The pitch link (fig. 2) is a vital element found in the helicopter made from aluminium alloy. The pitch variations for the tail rotor blades are transmitted with the help of a servo drive. The servo drive acts on the drive plate connected at the blade sleeves through the pitch links with the help of a mast.

The tension testing device (fig. 3) for the pitch links contains the following components:

- the pitch links' fastening device (3) that allows to fasten the pitch link (2) similar with the fastening on the helicopter;
- the pitch links tensioning device (1) with the help of an electric engine (4) through trapezoidal belts drives the cam mechanism, which at its turn through the kinematic chain determines an alternating movement of 2-5 cm for the links;
- with the help of an elastic element this alternating movement generates a dynamic force in the tested link;

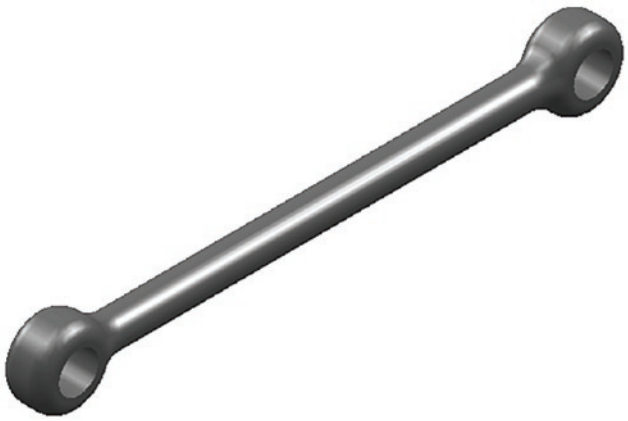


Fig. 2. Pitch link

- a cycle counter (5) indicates the number of cycles of fatigue testing of the pitch link;
- the control and automation (6) installation represents a part of the experimental stand that has the role of starting the electric engine and to stop the stand when a fissure appears, by decoupling the force and the engine.

3. The acceleration function method

The equivalence between the accelerated and normal testing requires that the following conditions are met [2]:

- the applied stress must not alter the physical mechanism for producing the faults: within the accelerated testing there must not be any other fault types or new faults related to the deteriorations, the only difference being the increase for of the occurrence of the fault.
- the conservation laws for the failing time distributions: the distribution functions have to remain the same for different defects, with the condition of increased speed in the appearance of the faults.

The first principle represents the basis for adopting the level of the applied overload, and the second principle certifies that the level of the overload does not exceed the allowable loading. The statistical analysis of the accelerated reliability testing is done in relation to the

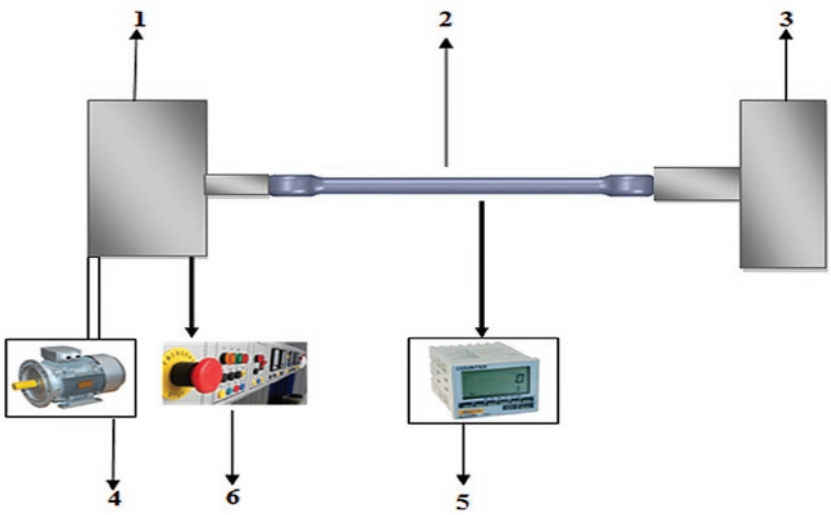


Fig. 3. Testing bench for the rotor pitch links

known rated normal level (from the producer's testing data sheet). The interpretation of the accelerated testing results is done in relation to the normal ones in known (standardized) normal conditions. It is necessary to know the rated stress reliability characteristic.

In this paper the accelerated testing will be validated by modelling the distribution laws (Weibull) of the failure times and the related parameters. When the modelling refers to the distribution law of the failing times, like the Weibull distribution, it is mandatory the values of the ß shape parameter from the normal testing (from the product data sheet) and from the accelerated testing are close, although the (η) scale parameter varies.

The most important aspect regarding these types of accelerated reliability tests is the equivalence between these and the normal level testing. The equivalence relationship can be obtained on the basis of equal reliability postulate which represents the theoretical foundation for accelerated testing. If for all the positive values of t the Rsn(t) > Rsa(t) is true then sn>sa. The equal reliability postulate shows the fact that for two stress levels (sn – normal and sa – accelerated) exists the equality of the reliability functions, as shown in the expression:

$$Rs_n(t) = Rs_a(\omega). \tag{1}$$

This equation (2) has a graphical correspondence in fig. 4 and implies a relationship between the moments (t, ω) through an acceleration function:

$$t = a(\omega). \tag{2}$$

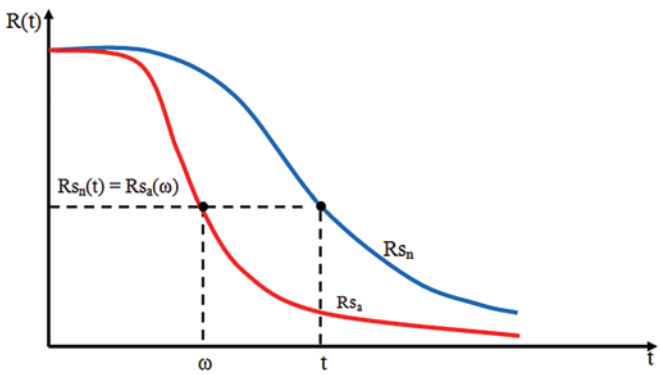


Fig. 4. Acceleration function

In the accelerated testing conditions, the accelerating function a(t) is a monotonically increasing function having the following properties:

$$\begin{cases} a(0) = 0 \\ \lim\limits_{\omega \to \infty} a(\omega) \to \infty \end{cases}. \tag{3}$$

The function a(ω) can be analytically expressed if the random variable's distribution law is known. In the case of the two parameter Weibull distribution we obtain the relationships:

$$Rs_n(t) = e^{-\left(\frac{t}{\eta_0}\right)^{\beta_0}}. \tag{4}$$

$$Rs_a(\omega) = e^{-\left(\frac{\omega}{\eta}\right)^{\beta}}. \tag{5}$$

From the equations (1), (4) and (5) we obtain the acceleration function relationship for the two parameter Weibull distribution:

$$t = \frac{\eta_0}{\eta^{\left(\frac{\beta}{\beta_0}\right)}} \cdot \omega^{\left(\frac{\beta}{\beta_0}\right)} = F \cdot \omega^n = a(\omega). \tag{6}$$

where the parameters F and n are constant (for given testing conditions) and can be experimentally determined. In this way, starting from the expression (6) the values for the variable parameters and for the normal loading distribution indicators can be obtained, if the values for the accelerated stresses are known.

4. Statistical analysis of accelerated reliability results

After the preceding steps of the accelerated reliability tests have been performed, the following statistical processing algorithm for the accelerated testing data was adopted, as represented in fig. 5.

This statistical data processing algorithm starts with the experimental data acquisition and finishes with determining the reliability indicators for the normal testing level. The results from the pitch link testing for the 3 accelerated testing levels are shown in Table 1.

Following the statistical analysis of the experimental data, for the three stress levels the most relevant distribution was the parameter Weibull distribution. In fig. 6 are represented the probability density functions for the three accelerated levels. To test the hypothesis which states that the distribution law for the number of cycles is Weibull type, one of the testing and certifying tests of the statistical hypothesis can be used. For the statistical data's distribution check obtained from the accelerated reliability testing, the Anderson-Darling test is used. This test compares the empirical cumulative distribution function of the sample data with the distribution expected if the data were normal. If the observed difference is adequately large, the null hypothesis of population normality will be rejected.

The most adequate acceleration model for data obtained from acceleration tests where the failure mode is through fatigue is the Inverse Power Law (IPL) – Weibull model [19]. In order to determine the number of cycles until failure and the reliability indicators for the

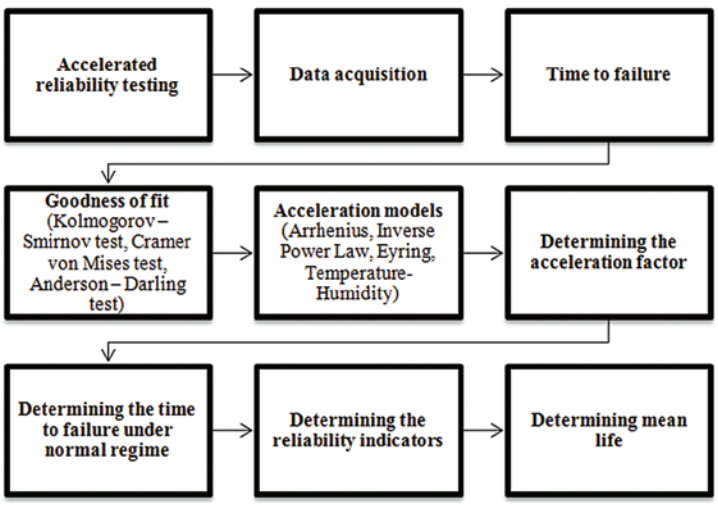


Fig. 5. The statistical processing algorithm for experimental data from accelerated reliability testing

Table 1. Accelerated reliability testing results of the rotor pitch links

No.	The number of cycles to failure in accelerated conditions	Tensile force [kN]
1	321518	17
2	384320	17
3	415470	17
4	423218	17
5	453514	17
6	214012	20
7	256080	20
8	276980	20
9	282012	20
10	302076	20
11	71437	22
12	85490	22
13	92327	22
14	94164	22
15	100891	22

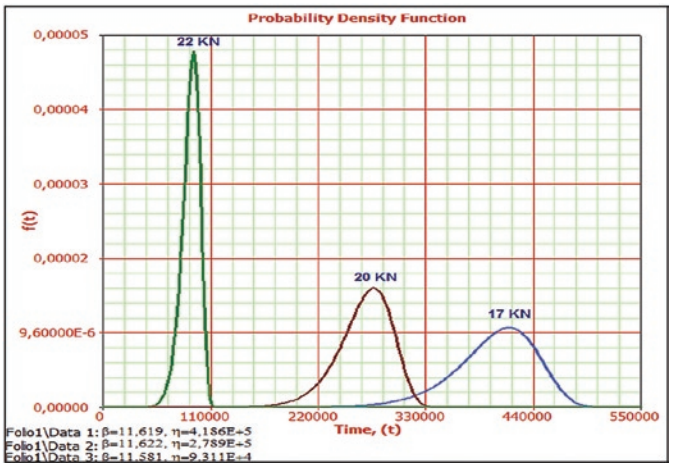


Fig. 6. Probability density plots

normal testing level (14 kN) for the pitch links, the data retrieved from the accelerated conditions were processed with the Weibull and ALTA 9 software. The three parameters characteristic to IPL-Weibull model are calculated by the maximum likelihood estimation method for the accelerated data, obtaining the following values: β=3.73; k=2.18E-12; n=4.86. The acceleration factor for the pitch links using the IPL-Weibull model has the following graphical representation (fig. 7). The acceleration factor is a number that describes a product's life at an accelerated stress level compared to the life at the normal stress level. The acceleration factor for tensile force of 22 KN is approximately 9.

As it can be seen in Table 2, the acceleration factor and standard deviation depend on the acceleration model relationship and are thus a function of tensile force.

The reliability indicators in normal conditions (F=15 kN) for the pitch links are determined using equations specific to the IPL – Weibull model according to the number of cycles until failure. In the fig. 8 there is a 3D representation of the reliability – time – tensile stresses. The reliability function represents an essential quantitative measure of reliability and has an important practical utility in the study of accelerated reliability tests. The reliability values depending

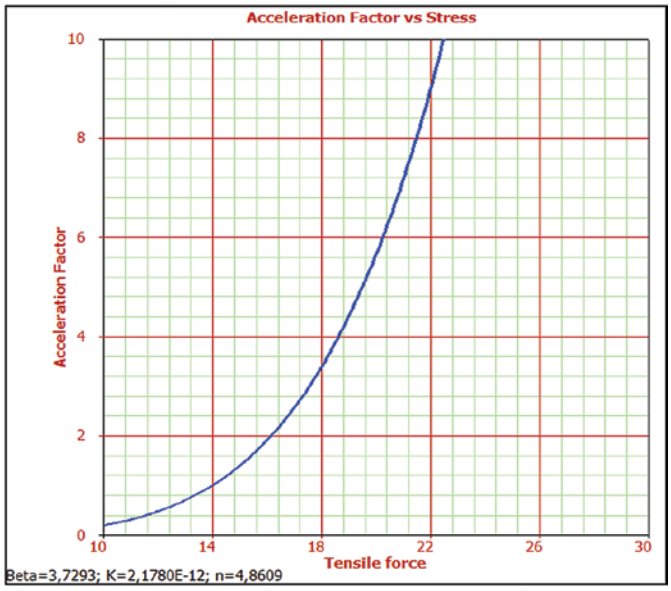


Fig. 7. The variation of the acceleration factor in relation to the tensile force

Table 2. The acceleration factor and standard deviation values in relation to the tensile force

Tensile Force [kN]	Acceleration factor	Standard deviation
14	1	332521
15	1.398	237777
16	1.914	173750
17	2.570	129402
18	3.393	98011
19	4.412	75359
20	5.662	58729
21	7.177	46329
22	8.998	36953

on the number of cycles to failure and on the stress level in normal testing (15 KN).

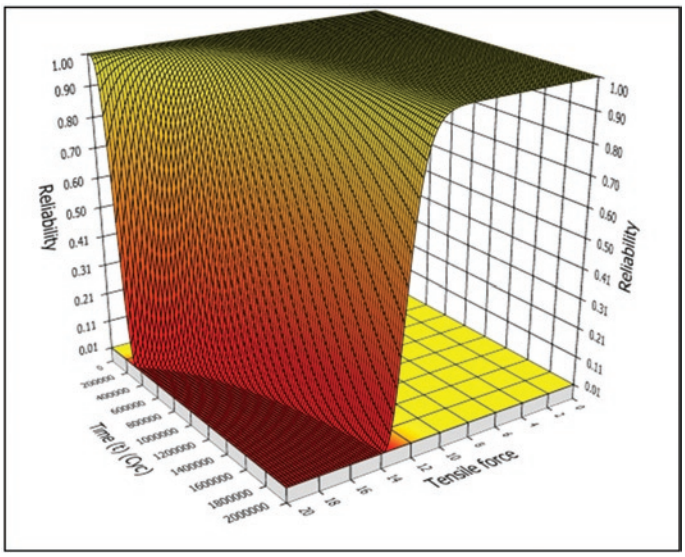
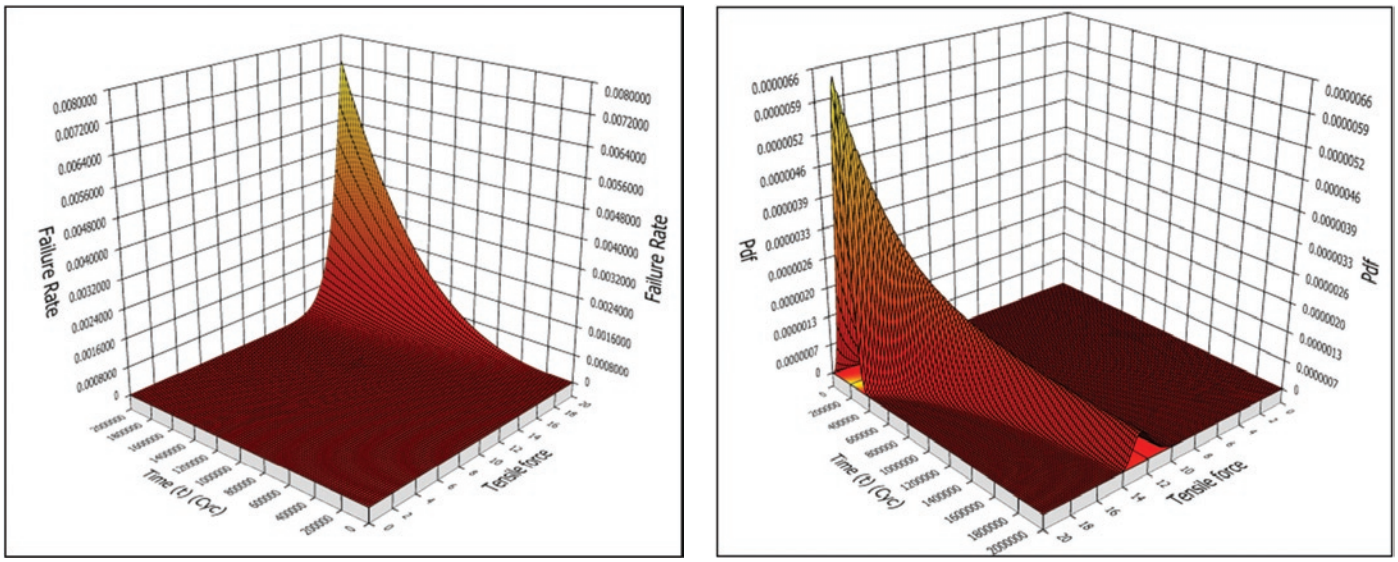


Fig. 8. Reliability function

Table 3. The dependence of the reliability indicators as function of the number of cycles to failure in normal conditions

The number of cycles to failure in normal conditions	Reliability R(t)	Unreliability F(t)	Probability density function $f(t)\cdot 10^{-6}$
642824	0.954	0.046	0.469
769280	0.890	0.110	0.704
826192	0.825	0.175	0.811
830803	0.760	0.240	0.820
847333	0.695	0.305	0.850
907866	0.630	0.370	0.954
987571	0.565	0.435	1.067
1067616	0.500	0.500	1.139
1087526	0.434	0.566	1.149
1165376	0.369	0.631	1.154
1211718	0.304	0.696	1.130
1449903	0.239	0.761	0.754
1568237	0.174	0.826	0.501
1596728	0.109	0.891	0.443
1710329	0.045	0.955	0.248



a) Failure rate *b) Probability density function*

Fig. 9. Reliability indicators

In fig. 9.a. is represented in 3D the failure rate – time – tensile stress. Probability density function 3D plot is represented in fig.9.b. for the pitch links.

The values for the main reliability indicators (reliability function, unreliability function and probability density function) are described in Table 3.

The fatigue life assessment for the helicopter pitch links represents one of the main objectives of this paper. In order to determine the life characteristic specific to the Weibull distribution, the graphical method is used. The mean number of cycles in normal conditions estimate the time at which 63.2% of the tested control rods are expected to fail for the three accelerated levels (17, 20 și 22 KN). At the intersection of the Eta curve, which estimates the mean (63.2%), with the axis from the normal stress level of 14 KN, finds itself the mean number of cycles to failure in normal conditions of the pitch links, which is 1232306 (fig. 10).

4. Validation of the accelerated reliability tests

The validation of the pitch links' accelerated reliability testing will be done as follows [15]:

- statistically it is done through the equivalence between the accelerated reliability testing and the normal testing conditions for the pitch links using graphically the reliability functions;
- also to certify the tests, it is mandatory that the values for the ß parameter in normal testing conditions and in the accelerated testing level are close.

The number of cycles until failure determined using the accelerated testing will be graphically represented, as well as the number of cycles until failure provided by the pitch links manufacturer in the normal testing conditions. As it can be seen from fig. 11 the β shape parameter for the number of cycles until failure determined using the accelerated testing has a value of 4.0523 and the corresponding pa-

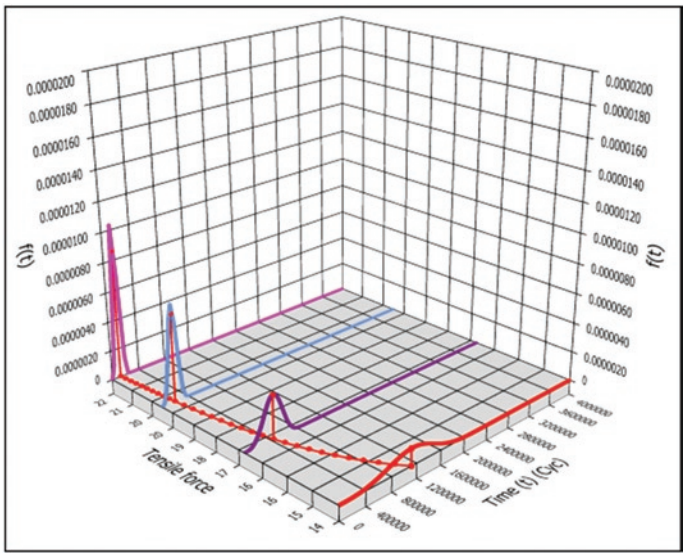


Fig. 10. Determining the mean number of cycles to failure of the pitch links in normal conditions

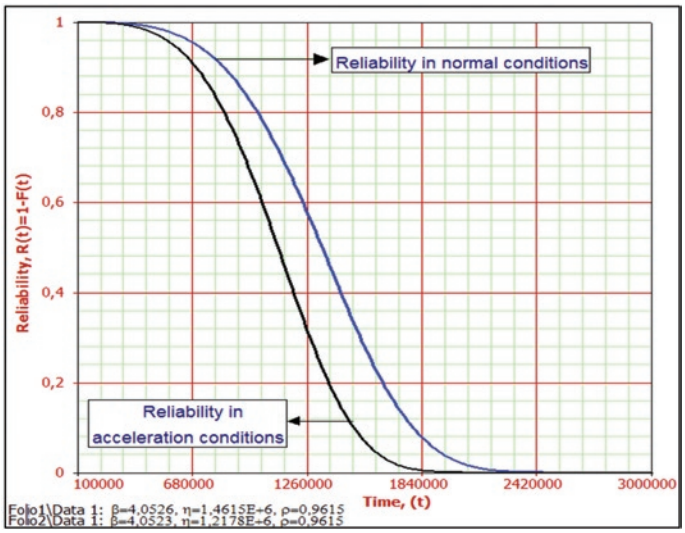


Fig. 11. The plot of reliability functions – validation of the accelerated reliability testing

rameter for the number of cycles until failure provided by the pitch links manufacturer has a value of 4.0526.

5. Conclusion

Usually, the analysis of the behavior of aerospace components during (normal) use is made based of their lifetimes, obtained by following their functioning in normal operation conditions. But in many situations, the determining of the life time of aerospace components in a reasonable timeframe is difficult or even impossible to accomplish, due to various reasons, such as: the very long life time of the products with high reliability, which in some cases can be in the order of years; the very short period of time between the design and the launch into fabrication; the continuous change of testing conditions where normal functioning regimes are used.

The main challenge that emerges by the reliability tests is their time span. To eliminate this shortcoming the accelerated reliability testing techniques are being implemented. Because the helicopter components are exposed simultaneously to various stresses in order to reduce the testing time there will be an amplification of those stresses that can determine the failure of the elements without modifying the degradation process.

In this paper, the validation of the accelerated tests was done by using the two principles: the physical mechanism of failure must not change and the conservation of failure time distribution law. By comparing the number of cycles in normal testing conditions (provided by the manufacturer) and the number of cycles that resulted from the accelerated testing, one can conclude that by using the accelerated reliability testing for the pitch links of the helicopter the number of cycles until failure was reduced by 4.5 times, a result that carries with it a major material cost reduction.

Acknowledgement

This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the project number POSDRU/89/1.5/S/59323. We hereby acknowledge the structural founds project PRO-DD (POS-CCE, O.2.2.1., ID 123, SMIS 2637, ctr. No 11/2009) for providing the infrastructure used in this work.

References

1. Asi O, Yeşil Ö. Failure analysis of an aircraft nose landing gear piston rod end. Engineering Failure Analysis 2013; 32: 283 – 291, https://doi.org/10.1016/j. engfailanal.2013.04.011.
2. Chang M-S, Lee C-S, Choi B-O, Kang B-S. Study on validation for accelerated life tests of pneumatic cylinders based on the test results of normal use conditions. Journal of Mechanical Science and Technology 2017; 31(6): 2739-2745, https://doi.org/10. 1007/s12206-017-0517-2.
3. Grbovic A, Rasuo B. FEM based fatigue crack growth predictions for spar of light aircraft under variable amplitude loading. Engineering Failure Analysis 2012; 26: 50 –64, https://doi.org/10.1016/j.engfailanal.2012.07.003.
4. IAR Ghimbav. Helicopter Flight Training Manual of IAR 330, 2000.
5. Kalaiselvan C, Rao L B. Accelerated life testing of nano ceramic capacitors and capacitor test boards using non-parametric method. Measurement 2016; 88: 58-65, https://doi.org/10.1016/j.measurement.2016.03.035.
6. Kececioglu D B. Reliability Engineering Handbook. New Jersey: PTR Prentice - Hall, Vol. I, 1991.
7. Kececioglu D B. Reliability & Life Testing Handbook. New Jersey: PTR Prentice-Hall, Vol. II, 1994.
8. Kim G-H, Lu H. Accelerated fatigue life testing of polycarbonate at low frequency under isothermal condition. Polymer Testing 2008; 27(1):114-121, https://doi.org/10. 1016 /j.polymertesting.2007.09.011.
9. Klyatis L M. Accelerated Reliability and Durability Testing Technology. New Jersey: Wiley, 2012.
10. Koo H-J, Kim Y-K. Reliability assessment of seat belt webbings through accelerated life testing. Polymer Testing 2005; 24(3): 309-315, https://doi.org/10.1016/j. polymertesting.2004.11.005.
11. Lewi E E. Introduction to Reliability Engineering. New Jersey: Wiley, 1995.
12. Lok S K, Paul J M, Upendranath V. Prescience Life of Landing Gear using Multiaxial Fatigue Numerical Analysis. Procedia Engineering

2014; 86: 775 – 779, https://doi. org /10.1016/j.proeng.2014.11.097.
13. Orkisz M, Święch Ł, Zacharzewski J. Fatigue tests of motor glider wing's composite spar. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14(3): 228- 232.
14. Özsoy S, Çelik M, Suat Kadıoğlu F. An accelerated life test approach for aerospace structural components. Engineering Failure Analysis 2008; 15(7): 946-957, https://doi. org/10.1016/j.engfailanal.2007.10.015.
15. Reliasoft. Accelerated life testing reference, Reliasoft publishing, 2009.
16. Shahani A R, Mohammadi S. Damage tolerance approach for analyzing a helicopter main rotor blade. Engineering Failure Analysis 2015; 57: 56-71, https://doi.org/ 10.1016/j.engfailanal.2015.07.025.
17. Tao S, Tan J, Haowen W. Investigation of rotor control system loads. Chinese Journal of Aeronautics 2013; 26(5): 1114-1124, https://doi.org/10.1016/j.cja.2013.07.029.
18. Van der Ven H, Bakker R J J, Van Tongeren J H, Bos M J, Münninghoff N. A modelling framework for the calculation of structural loads for fatigue life prediction of helicopter airframe components. Aerospace Science and Technology 2012; 23(1): 26-33, https://doi.org/10.1016/j.ast.2011.09.010.
19. Zaharia S M, Martinescu I. Reliability Testing. Brasov: Transilvania University Press Brasov, 2012.
20. Zaharia S M, Martinescu I. Management of accelerated reliability testing. Tehnički vjesnik 2016; 23(5): 1447-1455, https://doi.org/10.17559/TV-20141119153642.
21. Zaharia S M, Martinescu I, Morariu C O. Life time prediction using accelerated test data of the specimens from mechanical element. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14(2): 99 – 106.
22. Zhang C, Wang S, Bai G. An accelerated life test model for solid lubricated bearings based on dependence analysis and proportional hazard effect. Acta Astronautica 2014; 95: 30-36, https://doi.org/10.1016/j.actaastro.2013.10.019.
23. Zhang M, Meng Q, Hua W, Shi S, Hu M, Zhang X. Damage mechanics method for fatigue life prediction of Pitch-Change-Link. International Journal of Fatigue, 2010; 32(10): 1683-1688, https://doi.org/10.1016/j.ijfatigue.2010.04.001.

Sebastian Marian ZAHARIA
Department of Manufacturing Engineering
Transilvania University of Brasov
B-dul Eroilor nr. 29, 500036, Brasov, Romania

E-mail: zaharia_sebastian@unitbv.ro

Article citation info:
KAMIŃSKI Z, RADZAJEWSKI P . Calculations of the optimal distribution of brake force in agricultural vehicles categories R3 and R4. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 645–653, http://dx.doi.org/10.17531/ein.2019.4.14.

Zbigniew KAMIŃSKI
Paweł RADZAJEWSKI

CALCULATIONS OF THE OPTIMAL DISTRIBUTION OF BRAKE FORCE IN AGRICULTURAL VEHICLES CATEGORIES R3 AND R4

OBLICZENIA OPTYMALNEGO ROZDZIAŁU SIŁ HAMUJĄCYCH W PRZYCZEPACH ROLNICZYCH KATEGORII R3 I R4*

Fulfilling the requirements of the EU Directive 2015/68 in the area of braking for agricultural trailers depends on the proper selection of individual components of the braking system. This paper describes the requirements regarding braking performance and distribution of brake forces in agricultural trailers in R3 and R4 categories. On this basis, a methodology for calculating the optimal linear distribution of brake forces, characteristic for agricultural trailers with pneumatic braking systems, has been developed. The examples of calculation of an optimal distribution of brake forces for a two- and three-axle trailer with a tandem suspension system of the rear axle assembly have been provided. The optimization algorithm with the Monte Carlo method has been described, based on which a computer program was developed to select a linear distribution of brake forces in a three-axle trailer with 'walking beam' and 'bogie' suspensions. The presented calculations can be used in the design process to select the parameters of wheel braking mechanisms and then the characteristics of the pneumatic valves of the braking system.

***Keywords**: brake force distribution, optimization, agricultural vehicles, braking systems.*

Spełnienie wymagań Dyrektywy UE 2015/68 w zakresie hamowania przyczep rolniczych zależy od właściwego doboru poszczególnych komponentów układu hamulcowego. W pracy opisano wymagania dotyczące skuteczności hamowania oraz rozdziału sił hamujących w przyczepach rolniczych kategorii R3 i R4. Na tej podstawie opracowano metodykę obliczeń optymalnego liniowego rozdziału sił hamujących, charakterystycznego dla przyczep rolniczych z pneumatycznymi układami hamulcowymi. Zamieszczono przykłady obliczeń optymalnego rozdziału sił hamujących dla przyczepy dwu i trzyosiowej z tandemowym układem zawieszenia zespołu osi tylnych. Opisano algorytm optymalizacji metodą Monte Carlo, na podstawie którego opracowano program komputerowy do doboru liniowego rozdziału sił hamujących w przyczepie trzyosiowej z zawieszeniem „walking beam" i „bogie". Przedstawione obliczenia można wykorzystać w procesie projektowania do doboru parametrów kołowych mechanizmów hamulcowych, a następnie charakterystyk zaworów pneumatycznych układu hamulcowego.

***Słowa kluczowe**: rozkład siły hamowania, optymalizacja, pojazdy rolnicze, układy hamulcowe*

1. Introduction

In trailers and towed agricultural machines, pneumatic or hydraulic braking systems powered and controlled from an agricultural tractor are used most often [4, 8, 17, 26, 27, 28]. At present, inertial overrun brakes can only be used in low-speed towed vehicles (v≤40 km/h) with a total weight of less than 8000 kg and in high-speed vehicles (v>40 km/h) with a total weight not exceeding 3500 kg [2]. The service brakes of a tractor are driven by mechanical, hydraulic or air drive systems. The selection of drive system and energy source depends on the design and weight of the tractor.

Low and medium power tractors use simple and inexpensive hydraulic braking systems without power assistance [14]. Tractors with greater power use, first of all, hydraulic systems powered from the tractor hydraulic and pneumatic braking systems [17, 19, 27]. In low power tractors, mechanical brakes are still popular due to the costs.

The cooperation between tractor and trailer braking system is ensured by a trailer control valve (pneumatic or hydraulic) mounted in the tractor. Depending on the type of service brakes used in a tractor, the trailer control valves are mechanically, pneumatically or hydraulically actuated [12, 13, 22].

High-performance braking systems are a critical feature of modern agricultural vehicles. The new EU Regulation on Agricultural Vehicles [2], which has been in effect since 2016, includes a number of novel and much higher requirements for the braking performance of tractors and trailers, compatibility, safety and stability standards, including the introduction of ABS for vehicles traveling at speeds above 60 km/h.

For all categories of towed vehicles, the required braking rate has been increased. For vehicles with a total weight of over 3500 kg (category R3 and R4 agricultural trailers and towed agricultural machinery of category S2) and moving at a speed of over 40 km/h, a requirement of a specific distribution of brake forces between the axles of a vehicle has been introduced. As a result, it is possible to meet the requirement of achieving a sufficiently large relative deceleration (braking rate, i.e. the quotient of vehicle deceleration and the gravity $z=d/g$), conditioning the achievement of a short stopping distance and ensuring the directional stability of the braking vehicle in all traffic conditions. Similarly to the regulations concerning wheeled vehicles [24], no separate recommendations for vehicle combination were formulated, particular parts of tractor-trailer unit are treated as if they were single vehicles.

In order to adjust the distribution of braking forces between a tractor and a towed vehicle, compatibility requirements have been introduced for the first time in the form of permissible change areas

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

for the braking rate of a towing vehicle and a towed vehicle in a pressure function in the control line. The compliance with compatibility requirements as well as requirements regarding high-speed operation (response time of less than 0.6 s [2]), contribute to the shortening of braking distance of tractor-trailer units and the reduction of forces in the coupling during emergency braking [21].

The adoption of the new European legislation in field of agricultural vehicle places high demands on manufacturers of agricultural trailers, tractors and machinery in terms of braking systems [5]. The total fulfillment of the requirements regarding efficiency, stability and compatibility of brakes for agricultural trailers depends on the correct selection and calculation of individual components of the braking system (brake mechanisms and actuators, valves, pipes and braking force correctors), taking into account the construction parameters of trailers including axle systems and the type of suspension used [1, 11, 16].

Engineering calculations of braking systems of agricultural trailers are divided into design (synthesis) and testing (analysis). The purpose of design calculations is to determine the basic design parameters of braking systems and their components, taking into account the given operational characteristics. Design calculations include, among others:

- determining the permissible distribution of brake forces,
- selection of characteristics of brake force correctors,
- calculation of forces and torques of braking of wheels on individual axes for a given distribution of brake forces,
- calculation or selection of brake mechanisms,
- selection of actuators and calculation of application device,
- choice of braking system concept and selection of its components (valves, pipes, etc.).

The verification calculations are aimed at building and analyzing the characteristics of a considered system when its construction parameters are known. Such verification calculations include, but are not limited to:

- calculation of braking efficiency at the prescribed minimum pressure in the braking system,
- calculation of static characteristics, including the adhesion utilized by axle in a pressure function in a brake cylinder (braking rate) and checking the course of tractor and trailer braking rates in a pressure function at the coupling head of control line (of compatibility bands of braking system),
- calculations of dynamic characteristics to check the speed of operation (time of reaction) and synchrony of operation of individual circuits in a braking system.

This paper describes the methodology for the optimal selection of distribution of brake forces in agricultural trailers in R3 and R4 categories. The calculations of the distribution of brake forces is the basis for design calculations of vehicle braking systems, as they have a significant impact on the selection of the basic mechanisms and components of a braking system and the braking system efficiency [23]. An example calculation of a linear distribution of brake forces in a two-axle trailer and a three-axle trailer with a tandem suspension system of the rear axle assembly is provided here. An algorithm for the optimization of a linear distribution of brake forces in a three-axle trailer using the classic Monte Carlo method is described here.

2. Requirements concerning the efficiency, stability and compatibility of the braking systems of vehicles in R3, R4 and S2 categories

In the process of the selection of braking forces distribution between the axles of a trailer (towed machine) it is necessary to aim of an ideal distribution. Then, the rates f_i of adhesion utilized by all axles are the same throughout the braking process and, therefore, equal to the braking rate z of the vehicle:

$$\frac{T_1}{R_1} = \frac{T_2}{R_2} = \ldots = \frac{T_i}{R_i} = f_i = z \tag{1}$$

where: T_i - braking force of the wheels of the i-th axle, R_i - normal reactions of the road surface on the wheels i-th axle, z - braking rate of the vehicle $z=\sum T_i/\sum R_i$

This distribution of brake forces is considered to be optimal because, with homogeneous surfaces, the achievement of the highest possible braking intensity in given conditions and the fulfillment of braking efficiency requirements with reserve are achieved (Table 1).

Table 1. The required braking efficiency of agricultural trailer's service brakes [2]

Vehicle category	Braking rate z [%] in p=6,5 bar	
	v≥30 km/h	v>30 km/h
Trailers R2, R3, R4	35%	50%
Towed machines S2	35%	50%

Due to the variable loading levels of trailers, it is practically impossible to achieve an ideal distribution of brake forces, even when using braking force regulators. Therefore, for high-speed agricultural vehicles (speed above 40 km/h), the allowable limits for derogation of adhesion utilization rates f_i for individual axles against the ideal distribution have been determined. From 2016, two solutions have been allowed, as shown in Figure 1 [2].

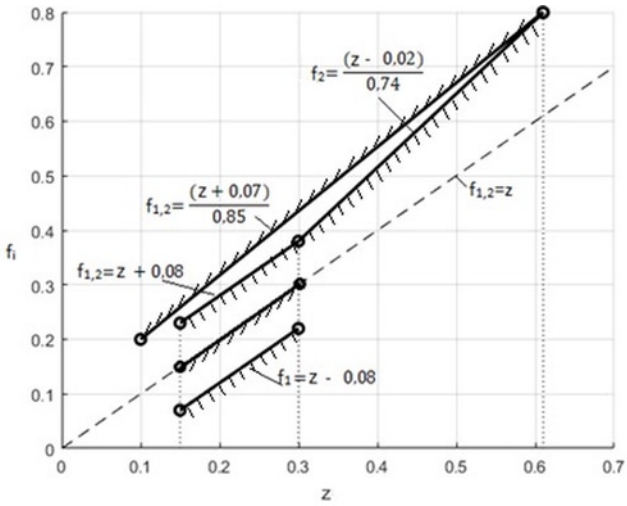


Fig.1. Limit values of adhesion utilization for both solutions

The first solution: the adhesion utilization rate for each axle must meet the condition of ensuring the minimum required braking performance:

$$\left.\begin{matrix} f_1 \\ f_2 \end{matrix}\right\} \le \frac{z+0.07}{0.85} \quad \text{for} \quad z = 0.1 \div 0.61 \tag{2}$$

and the condition of previous locking of the front axle wheels to ensure directional stability:

$$f_1 > z > f_2 \quad \text{for} \quad z = 0.15 \div 0.30 \tag{3}$$

The second solution: the adhesion utilization rates by the two axles should be within a given band, and then the limits of wheel locking are determined by the following relationships:

$$\begin{aligned} f_1 &\geq z - 0.08 \\ f_{1,2} &\leq z + 0.08 \end{aligned} \quad \text{for} \quad z = 0.15 \div 0.30 \tag{4}$$

In addition, the adhesion utilization curve for the rear axle should fulfill the condition:

$$f_2 \leq \frac{z - 0.02}{0.74} \quad \text{for} \quad z \geq 0.3 \tag{5}$$

The requirements described above also apply to trailers with more than two axles. Then, the adhesion utilization rates used by the front axle assembly and the rear axle assembly are calculated based on the relationship:

$$f_1 = \frac{\sum f_{1i} R_{1i}}{\sum R_{1i}} \qquad f_2 = \frac{\sum f_{2i} R_{2i}}{\sum R_{2i}} \tag{6}$$

The wheel locking requirements may be considered fulfilled if; for braking efficiency rates between 0.15 and 0.30, the adhesion utilized by at least one of the front axles is greater than that applied by at least one of the rear axles [2]:

$$f_{1i} > f_{2i} \quad \text{for any } i \tag{7}$$

In the considerations regarding the distribution of brake forces (2)-(5), each part of a road unit is treated as a single vehicle, without taking into account the braking control of towed vehicles. Hence, in order to ensure the compatibility of brake forces in a vehicle combination, acceptable bands of changes of braking ratios of individual vehicles for their characteristic load states in a control pressure function on the coupling head [2] have been determined - Fig. 2.

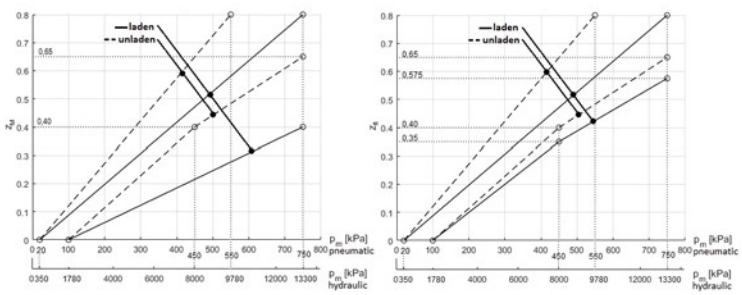


Fig. 2. Permissible bands of braking rate for tractors z_M and trailers z_R in a pressure function p_m in a control line

3. The determination of the permissible distribution area of brake forces in two-axle trailers

As shown in Figure 3, dynamic normal force on front and rear axles of a trailer on the horizontal road surface varies depending on the braking intensity (braking efficiency rate z) as follows:

$$R_1 = \frac{G}{L}(b + h \cdot z) \qquad R_2 = \frac{G}{L}(L - b - h \cdot z) \tag{8}$$

where: L – wheelbase of a trailer, h - height of the center of gravity, b – horizontal distance between center of gravity and rear axle.

The relative (relating to the G weight of a trailer) of braking force of the front axle γ_1 and rear γ_2 is calculated based on the relationship:

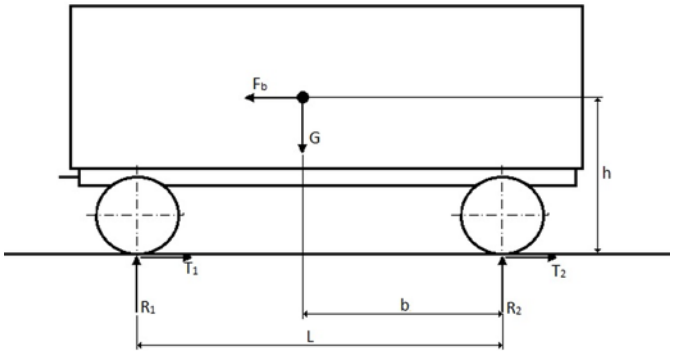


Fig. 3. Diagram showing forces acting upon a two-axle trailer during braking.

$$\gamma_1 = \frac{T_1}{G} = \frac{R_1 f_1}{G} = \left(\frac{b}{L} + \frac{h}{L} z\right) f_1 \qquad \gamma_2 = \frac{T_2}{G} = \frac{R_2 f_2}{G} = \left(1 - \frac{b}{L} - \frac{h}{L} z\right) f_2 \tag{9}$$

where: T_1, T_2 - braking forces of the front and rear axles, R_1, R_2 - normal reactions of the road surface on the wheels of i-th axle:

Under ideal braking conditions, the adhesion utilization ratios used by the front and rear axle of a trailer are indistinguishable and equal to the braking intensity f_1=f_2=z, and the distribution of brake forces is described by the parametric equation:

$$\gamma_1 = \left(\frac{b}{L} + \frac{h}{L} z\right) z \qquad \gamma_2 = \left(1 - \frac{b}{L} - \frac{h}{L} z\right) z \tag{10}$$

Using the dependences on the limit values of the adhesion rates used by the axles (2,3,4,5) along with the technical data of a trailer, it is possible to determine the lower and upper limit of the permissible distribution of brake forces in the diagram of relative of brake forces γ_2=f(γ_1). Graphic interpretation of the described recommendations according to the first solution is illustrated by lines AB and CD in Fig.4-a and Fig.5-a. The corresponding limitations of the brake forces in the coordinate system γ_1-γ_2 for an exemplary trailer when unladen and when laden are shown in Fig.4-b and Fig.5-b. The margin curves are calculated by substituting the adhesion utilization rates f_1, f_2 determined from the conditions (2), (3) to the relation (9).

In the second solution, the limitations of the acceptable area of adhesion utilization rates are marked by lines MN and JKL in Fig.4-c and Fig.5-c. The corresponding areas of relative brake forces according to the second solution are shown in Fig. 4-d for an unladen trailer, and for a laden trailer in Fig.5-d. Due to the restrictive nature of the condition (4) for the upper margin K'L' on the chart γ_2=f(γ_1), its scope was limited to the range of z=0.3÷0.61.

The equations of individual lines and margin curves in the $f_{1,2}$-z and γ_1-γ_2 systems along with the coordinates of individual points are summarized in Tab.2 and Tab.3.

4. The selection of linear distribution of brake forces in two-axle trailers

In air braking systems of agricultural trailers, braking force correctors with radial (linear) characteristics are usually used [8, 17, 28]. This characteristic is described by the equation of a straight line passing through the beginning of the coordinate system and a second selected point on the graph of relative of brake forces γ_2=f(γ_1) taking into account the area limitations described in the previous chapter. The procedure for determining the acceptable range of changes in the

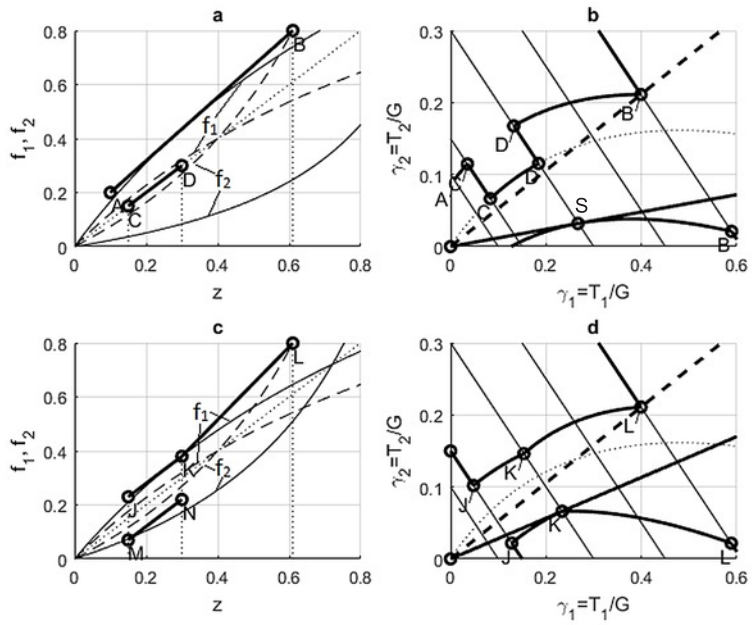


Fig. 4. Determining the parameters of constant distribution of brake forces for an unladen trailer weighing 4200 kg: a, c - runs of adhesion utilization rates used by axles f_1, f_2; b - boundary values of the distribution coefficient according to solution 1; d - boundary values of the distribution coefficient according to solution 2; L = 2.95 m; b = 1.47m; h = 1.15 m

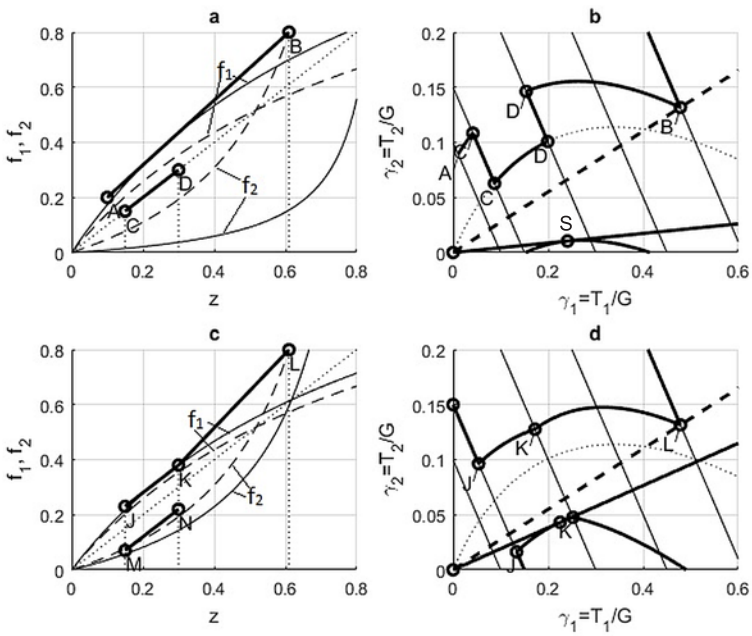


Fig.5. Determining the parameters of constant distribution of brake forces for a loaded trailer with a mass of 16250 kg: a, c - runs of adhesion utilization ratios used by axles f1, f2; b - limit values of the separation factor according to solution 1; d - limit values of the separation factor according to solution 2; L = 2.95 m; b = 1.47m; h = 1.63 m

directional coefficient $i_p=T_2/T_1=\gamma_2/\gamma_1$ of straight lines showing a constant distribution of brake forces should be carried out for laden and unladen vehicles.

In the first solution, the area of permissible linear distribution of brake forces is determined from the bottom; by a straight line OS tangent to the margin curve AB at point S (Fig.4-b, Fig.5-b), and from the top; by a straight line passing through point D or B' (select a straight line with a smaller value of a directional coefficient).

When using the second solution, the lower margin of the acceptable area is determined by a straight tangent line at the point T with the JK curve (Fig. 4-d). If the point of contact T lies outside the JK section of the margin curve then the direction coefficient of the margin line is determined on the basis of the coordinates of point K (Fig.5-d). The upper margin of the linear distribution of brake forces is determined by the straight line passing through point L' (Fig. 4-d, Fig. 5-d). Sometimes it can also be point N.

Since the straight line of the linear distribution of brake forces passes through the beginning of the coordinate system $\gamma_2=f(\gamma_1)$, its direction coefficient is calculated in each case from the ratio of the ordinate to the abscissa of the given characteristic point P:

$$i_P=\frac{\gamma_{2p}}{\gamma_{1p}}=\frac{1-b/L-z_p\cdot h/L}{b/L+z_p\cdot h/L} \tag{11}$$

Where: *P* - symbol of a characteristic point.

Using the dependence $z=\gamma_1+\gamma_2$, it is possible to describe the distribution of brake forces of individual axles for a given line by means of a parametric equation:

$$T_1=\frac{1}{1+i_P}G\cdot z \qquad T_2=\frac{i_P}{1+i_P}G\cdot z \tag{12}$$

in which the parameter is the braking rate *z*. The adhesion utilized rates by the axles on a given distribution line of braking force are calculated as follows:

$$f_1=\frac{T_1}{R_1}=\frac{z}{\left(\frac{b}{L}+z\frac{h}{L}\right)(1+i_P)} \qquad f_2=\frac{T_2}{R_2}=\frac{i_p\cdot z}{\left(1-\frac{b}{L}-z\frac{h}{L}\right)(1+i_p)} \tag{13}$$

The results of the calculation of the limit values of directional coefficients i_P for the considered two-axle trailer are summarized in Table 4.

The courses of coefficients f_1, f_2 of adhesion utilized by the axles, corresponding to the individual margin lines of the distribution of brake forces (tab. 4), calculated for dependence (13) for the 1st and 2nd solution, are shown in Fig. 4,5-a, c.

When choosing the range of changes of the real distribution of brake forces $i=T_2/T_1$ for an unladen and a laden trailer, one should strive to approximate with the upper straight margin line. This ensures a short braking distance with the simultaneous danger of the earlier locking of rear wheels in an intensity range greater than the braking rate at the intersection of the ideal distribution curve (10) with a constant distribution line.

The coefficient of adhesion utilization (the factor of using the vehicle's weight for braking) is a measure of the efficiency of a well-chosen distribution of brake forces of the vehicle for different values of μ:

$$\zeta(\mu)=\frac{T}{\mu\cdot G}=\frac{z}{\mu} \tag{14}$$

where: μ- coefficient of adhesion

In the search for an optimal value of a linear coefficient of a braking force distribution, the criterion of equality of minimum coefficient

Table 2. The requirements of braking efficiency and stability for trailers

Curve	Coordinate system $f_{1,2}$-z	Coordinate system γ_1-γ_2	Range
A-B	$z \geq 0.85 \cdot f_{1,2} - 0.07$	$\gamma_1 = \left(\frac{b}{L} + \frac{h}{L} z\right)\left(\frac{z+0.07}{0.85}\right)$	z=0.1-0.61
C-D	$z \leq f_{1,2}$	$\gamma_1 = \left(\frac{b}{L} + \frac{h}{L} z\right) z$	z=0.15-0.30
A'-C' D'-B'	$z \geq 0.85 \cdot f_{1,2} - 0.07$	$\gamma_1 = z - \left(1 - \frac{b}{L} - \frac{h}{L} z\right)\left(\frac{z+0.07}{0.85}\right)$	z=0.1-0,15 z=0.3-0.61
J-K	$z \leq f_{1,2} - 0.08$	$\gamma_1 = \min\begin{cases} \left(\frac{b}{L} + \frac{h}{L} z\right)(z+0.08) \\ z - \left(1 - \frac{b}{L} - \frac{h}{L} z\right)(z-0.08) \end{cases}$	z=0.15-0.30
M-N	$z \leq f_{1,2} + 0.08$	$\gamma_1 = \max\begin{cases} \left(\frac{b}{L} + \frac{h}{L} z\right)(z-0.08) \\ z - \left(1 - \frac{b}{L} - \frac{h}{L} z\right)(z+0.08) \end{cases}$	z=0.15-0.30
K-L	$z \geq 0.3 + 0.74\left(f_{1,2} - 0.38\right)$	$\gamma_1 = \left(\frac{b}{L} + \frac{h}{L} z\right)\left(\frac{z-0.3}{0.74} + 0.38\right)$	z=0.30-0.61
K'-L'	$z \geq 0.3 + 0.74\left(f_{1,2} - 0.38\right)$	$\gamma_1 = z - \left(1 - \frac{b}{L} - \frac{h}{L} z\right)\left(\frac{z-0.3}{0.74} + 0.38\right)$	z=0.30-0.61

Table 3. The coordinates of characteristic points; there is a dependency of $\gamma_2 = z - \gamma_1$ for all ranges.

Point	z	$f_{1,2}$	γ_1
A	0.10	0.20	$0.2(b/L + 0.1 \cdot h/L)$
B	0.61	0.80	$0.8(b/L + 0.61 \cdot h/L)$
C	0.15	0.15	$0.15(b/L + 0.15 \cdot h/L)$
D	0.30	0.30	$0.3(b/L + 0.3 \cdot h/L)$
A'	0.10	0.20	$0.1 - 0.2(1 - b/L - 0.1 \cdot h/L)$
C'	0.15	0.259	$0.15 - (0.22/0.85)(1 - b/L - 0.15 \cdot h/L)$
D'	0.30	0.435	$0.3 - (0.37/0.85)(1 - b/L - 0.3 \cdot h/L)$
B'	0.61	0.8	$0.61 - (0.68/0.85)(1 - b/L - 0.61 \cdot h/L)$
J	0.15	0.23	$\min\begin{cases} 0.23(b/L + 0.15 \cdot h/L) \\ 0.15 - 0.07(1 - b/L - 0.15 \cdot h/L) \end{cases}$
K	0.30	0.38	$\min\begin{cases} 0.38(b/L + 0.3 \cdot h/L) \\ 0.3 - 0.22(1 - b/L - 0.3 \cdot h/L) \end{cases}$
L	0.61	0.8	$(b/L + 0.61 \cdot h/L)(0.31/0.74 + 0.38)$
M	0.15	0.07	$\max\begin{cases} 0.07(b/L + 0.15 \cdot h/L) \\ 0.15 - 0.23(1 - b/L - 0.15 \cdot h/L) \end{cases}$
N	0.30	0.22	$\min\begin{cases} 0.22(b/L + 0.3 \cdot h/L) \\ 0.3 - 0.38(1 - b/L - 0.3 \cdot h/L) \end{cases}$
K'	0.30	0.38	$0.38(1 - b/L - 0.3 \cdot h/L)$
L'	0.61	0.8	$0.61 - (1 - b/L - 0.61 \cdot h/L)(0.31/0.74 + 0.38)$

Table 4. Limit values of directional coefficients for a linear distribution of brake forces for a two-axle trailer

Variant of the solution	An unladen trailer	A laden trailer
The first solution acc.to (2), (3)	$i_{min}=i_S=0.1202$	$i_{min}=i_S=0.0434$
	$i_{max}=i_{B'}=0.5293$	$i_{max}=i_{B'}=0.2754$
The second solution acc.to (4), (5)	$i_{min}=i_K=0.2832$	$i_{min}=i_T=0.1914$
	$i_{max}=i_{L'}=0.5383$	$i_{max}=i_{L'}=0.2750$
Optimal coefficient acc.to (16)	$_{iopt}=0.5759$	$_{iopt}=0.4463$

of adhesion utilization for two extreme values of adhesion coefficients $\mu_1<\mu<\mu_2$ [6] characterizing the vehicle operation conditions is used:

$$\zeta(\mu_1)=\zeta(\mu_2) \tag{15}$$

or the criterion of maximizing the average adhesion utilization coefficient $\zeta(\mu)$ in a given range (μ_1, μ_2) [7]:

$$\zeta_{sr}=\frac{1}{\mu_2-\mu_1}\int_{\mu_1}^{\mu_2}\zeta(\mu)d\mu \tag{16}$$

In the case of two-axle trailers, the same optimal value of the coefficient of adhesion is obtained for both criteria [10]:

$$\mu_{op}=\mu_1+\frac{b}{L}(\mu_2-\mu_1) \tag{17}$$

On this basis, it is easy to determine the optimal value of the directional coefficient for the distribution of brake forces [9, 10]:

$$i_{opt}=\frac{1-b/L-\mu_{op}\cdot h/L}{b/L+\mu_{op}\cdot h/L} \tag{18}$$

by changing b/L and h/L values respectively for an unladen and a laden trailer. In calculations for agricultural trailers, one can take μ_1=0.2 and μ_2=0.5 [9]. The optimal line of constant distribution of brake forces calculated from the formula (18) must lie within the permissible range defined by simple margin lines (Tab.4). In the case under consideration here, the optimum values of the directional coefficient for an unladen and a laden trailer are greater than the maximum permissible value. Nevertheless, this fact supports the adoption of higher values of braking distribution coefficients, close to the optimal values (the second solution).

If the lines of distribution of relative of brake forces pass through point B' (the first solution) or L' (the second solution), the distribution coefficients are identical. In both cases, the curve $f_2(z)$ of adhesion utilized by the rear axle passes through the point of coordinates z=0.61 and $f_2(0.61)$=0.8. Calculating the brake forces for this point:

$$\begin{aligned} T_2&=f_2R_2=0.8G\left(1-\frac{b}{L}-0.61\frac{h}{L}\right)\\ T_1&=G\cdot z-T_2=G\left[0.61-0.8\left(1-\frac{b}{L}-0.61\frac{h}{L}\right)\right]\end{aligned} \tag{19}$$

the following statement is given for the factor of distribution of brake forces:

$$i_P=\frac{T_2}{T_1}=\frac{0.8(1-b/L-0.61h/L)}{0.61-0.8(1-b/L-0.61h/L)} \tag{20}$$

The differences between maximum values i_{max} given in Tab. 4 for both solutions result from rounding of inequality coefficients (5). For precise calculations, the divisor in the statement (5) should be 0.7381.

5. The selection of linear distribution of brake forces in three-axle trailers

In agricultural three-axle trailers, two rear axles are located close and work in a tandem arrangement. The braking force on the rear axle assembly must be distributed according to the load distribution between the tandem suspension axles. The system of forces acting on an agricultural three-axle trailer with tandem suspension of 'walking beam' type is shown in Fig.6. In the adopted calculation model, it is assumed that the un-sprung weight of the tandem axles will be omitted, which means that the forces of gravity and the inertia of the suspension are omitted.

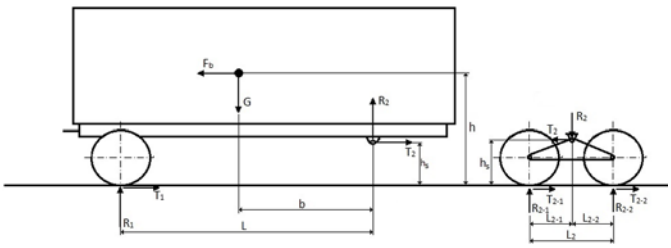


Fig. 6. Scheme of forces acting on a three-axle trailer with a 'walking beam' tandem suspension

The balance of forces and moments acting on the trailer takes the form:

$$\sum F_x=T_1+T_2-G\cdot z=0 \tag{21}$$

$$\sum F_y=R_1+R_2-G=0 \tag{22}$$

$$\sum M_2=G\cdot b+G\cdot z(h-h_s)+T_1\cdot h_s-R_1\cdot L=0 \tag{23}$$

After solving the above system of equations, the relationships describing reactions R_1 and R_2 acting on the trailer are obtained:

$$R_1=G\left(\frac{b}{L}+z\frac{h-h_s}{L}\right)+T_1\frac{h_s}{L} \quad \text{or} \quad R_1=G\left(\frac{b}{L}+z\frac{h}{L}\right)-T_2\frac{h_s}{L} \tag{24}$$

$$R_2=G\left(1-\frac{b}{L}-z\frac{h-h_s}{L}\right)-T_1\frac{h_s}{L} \quad \text{or} \quad R_2=G\left(1-\frac{b}{L}-z\frac{h}{L}\right)+T_2\frac{h_s}{L} \tag{25}$$

where:

$$T_1=\frac{1}{1+i_P}G\cdot z \qquad T_2=G\cdot z-T_1 \tag{26}$$

The balance of forces and moments acting on the tandem suspension takes the form:

$$\sum F_x=T_{21}+T_{22}-T_2=0 \tag{27}$$

$$\sum F_y = R_{21} + R_{22} - R_2 = 0 \tag{28}$$

$$\sum M_{21} = T_2 \cdot h_s + R_{22} L_2 - R_2 \cdot L_{21} = 0 \tag{29}$$

In order to determine the distribution of the braking force T2 between the wheels of the tandem axle unit, a linear distribution of brake forces is assumed:

$$\frac{T_{22}}{T_{21}} = i_S$$

By solving a system of equations (27)-(29), we obtain the value of forces acting on wheels in tandem axles during braking:

$$R_{21} = R_2 \frac{L_{22}}{L_2} + T_2 \frac{h_s}{L_2} \qquad R_{22} = R_2 \frac{L_{21}}{L_2} - T_2 \frac{h_s}{L_2} \tag{31}$$

$$T_{21} = \frac{T_2}{1 + i_S} \qquad T_{22} = \frac{i_S}{1 + i_S} T_2 \tag{32}$$

Formulas (24), (25) and (31), (32) can also be used for a 'boogie' tandem suspension.

The Monte Carlo method [3], [18], [20] was used to search for an acceptable range of variability of i_P and i_S coefficients of distribution of brake forces in order to find optimal solutions. A block diagram of an algorithm for an optimal selection of the braking force distribution coefficients is shown in Fig.7. On its basis, a computer program was developed in the Matlab [25] environment.

The optimum values of the braking force distribution coefficients were determined in the process of minimizing the objective function in the form of:

$$FC = \frac{w_1 (f_1 - f_2)^2 + w_2 (f_{21} - f_{22})^2}{w_1 + w_2} \tag{33}$$

where: w_i - weighting factors.

The function formulated this way prefers the solutions that approximate the adhesion utilized f_i by individual axles.

Before calculating the objective function, we have checked the inequality constraints (4), (5) for the second solution:

$$\begin{aligned} f_1 &\ge f_1^{down} = z - 0.08 \\ f_1 &\le f_1^{up} = z + 0.08 \end{aligned} \quad \text{for} \quad z = 0.15 \div 0.30$$

$$f_2 \le f_2^{up} = (z + 0.08)(0.15 \le z \le 0.30) + \left(\frac{z - 0.3}{0.7381} + 0.38 \right)(z \ge 0.30) \tag{34}$$

and condition (7):

$$f_1 > f_{21} \quad \text{or} \quad f_1 > f_{22} \quad \text{for} \quad z = 0.15 \div 0.30 \tag{35}$$

In addition, an extra condition has been adopted for the adhesion utilized rates of rear axle:

$$f_{2i} \le f_2^{up} \tag{36}$$

limiting excessive increase of coefficient f_{22} for $z \le 0.61$.

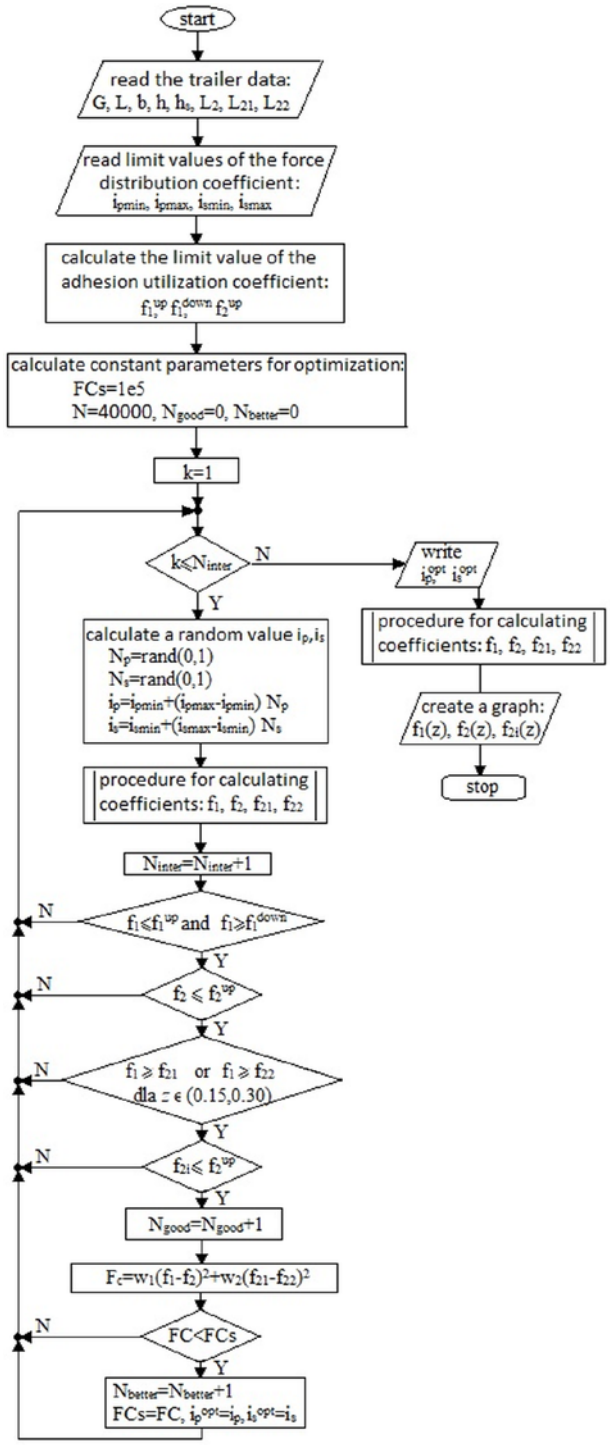


Fig. 7. A block diagram of an algorithm for the optimization of brake forces of a three-axle trailer using the Monte Carlo method (FC_s - initial value of the objective function, N - number of draws, N_{good} – a number of good solutions, meeting inequality constraints, N_{better} - number of better solutions, reducing the value of the objective function).

The results of the calculation of the distribution of braking force for an empty and a loaded trailer after several program start-ups are presented in Tab.5. The adopted number of draws was $N = 40{,}000$, w_1=0.6, w_2=0.4.

An example of the course of the adhesion utilization rates $f_i(z)$ through the axles for an optimal distribution of brake forces for an unladen and a laden trailer is shown in Fig.8.

Table 5. The results of the optimization of distribution of brake forces in a three-axle trailer

No.	1	2	3	4	5	Average
	An unladen trailer					
i_S	1.2971	1.2963	1.2944	1.2991	1.2955	1.2965
i_P	0.5150	0.5162	0.5155	0.5153	0.5157	0.5155
FC	1.7961	0.8805	0.4060	0.7386	1.5802	1.0803
	A laden trailer					
i_S	0.9753	0.9737	0.9749	0.9729	0.9725	0.9739
i_P	0.5326	0.5337	0.5314	0.5333	0.5331	0.5328
FC	0.5287	0.5616	0.3032	1.1820	0.6626	0.6476

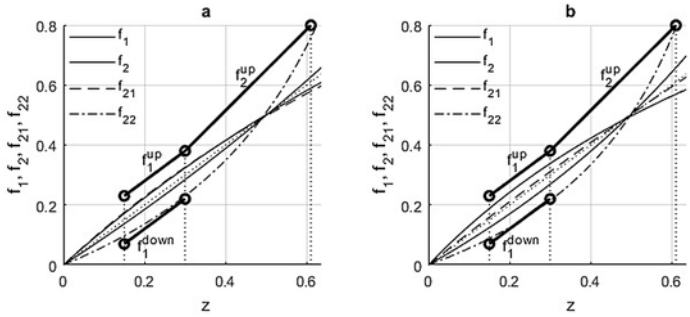


Fig. 8. The runs fi(z) for an optimal distribution of brake forces: a- for an empty trailer (is = 1.2994, ip = 0.5155, m = 7700kg, L = 5.15m, b = 1.85m, h = 1.8m, L2 = 1.345m, L21 = 0.72m, L22 = 0.62m, hs = 0.545m), b - for a loaded trailer (is = 0.9749 , ip = 0.5333, m = 24000kg, L = 5.15m, b = 1.85m, h = 1.8m, L2 = 1.36m, L21 = 0.73m, L22 = 0.63m, hs = 0.514m)

6. The calculations and selection of a brake mechanism and a starting mechanism

Knowing the values of the brake forces of individual axles, the design parameters of their braking mechanisms and starting mechanisms can be calculated. Here you can use the following dependence on the braking force of *i-th* axle [2]:

$$T_i = k \cdot \left(C - C_0\right) \cdot \eta \cdot BF / r_d + f_r R_i \tag{37}$$

where: k – a number of actuators per axle, C- the torque on the cam shaft generated by the actuator, C_0 – the threshold torque of the cam shaft necessary to create a measurable braking torque, η - mechanical efficiency, r_d - dynamic wheel radius, f_r – coefficients of the resistance of wheel rolling, R_i - load on the wheels of the *i-th* axle, BF - 'Brake factor' coefficient, is defined as follows [1]:

$$BF = \frac{C^* \cdot r_e}{2 r_b} \tag{38}$$

where: C^*- coefficient of efficiency (internal brake ratio) of the brake mechanism [15], r_e – effective friction radius, r_b - effective radius of an S-cam.

The torque on the cam shaft is the product of the Th_a force generated by the pneumatic actuator acting on the lever span with length l:

$$C = Th_a l \tag{39}$$

Using the experimental data of the manufacturers of actuators, one may express the useful force on the piston rod by means of:

$$Th_a = A \cdot p - B \tag{40}$$

where: A, B - experimental coefficients, p - pressure in the actuator chamber.

7. Summary and conclusions

The calculations described in this paper will enable the selection of an optimal linear distribution of brake forces in the design process of air braking systems on two- and three-axle agricultural trailers, in which the braking force correctors with radial characteristics are used. The calculations of the distribution of brake forces took into account the requirements of the EU Directive 2015/68 [2] in terms of braking efficiency and stability. The calculations of the distribution of brake forces form the basis of design calculations and enable, in the next stage of the design process, the selection of braking axle parameters (braking mechanism, actuator and applying device) and characteristics of brake valves.

The calculation of the distribution of brake forces made for a two-axle trailer with a capacity of about 16 tons, carried out for two possible variants of the distribution of brake forces, support the use of the second solution in the design calculations (according to requirements (4), (5)). The adhesion utilization rates calculated for this solution used by the axles are more similar to the straight line illustrating an ideal distribution of brake forces, in which the coefficients of adhesion utilized by each axle are the same and equal to the braking rate.

In the case of a two-axle trailer, the range of permissible changes of the linear coefficient of distribution of brake forces and its optimal value for various loading conditions are determined analytically, based on the graph of relative brake forces γ_2 (γ_1). However, in the case of three-axle trailers, in which the brake forces must be divided between the axles of the tandem assembly, faster results are obtained using optimization methods.

The developed algorithm looks for an optimal distribution of brake forces with the Monte Carlo method for trailers with 'walking beam' or 'bogie' type of rear axle suspensions, which can be easily adapted to the selection of brake forces distribution in trailers with other types of tandem axles by changing the block of procedure of calculation the adhesion utilized through axles (other dependencies on wheel reactions).

On the basis of the presented methodology, it is possible to develop rules for the distribution of brake forces of a trailer using braking force correctors with different characteristics than radial (linear) ones.

Acknowledgement

Project financed with the program of the Minister of Science and Higher Education under the name "Regionalna Inicjatywa Doskonałości" in the years 2019 - 2022 project number 011 / RID / 2018/19 amount of financing 12,000,000 PLN

Ministerstwo Nauki i Szkolnictwa Wyższego

References

1. Andrew J D. Braking of Road Vehicles. Oxford: Butterworth-Heinemann, 2014.
2. Commission Delegated Regulation (EU) 2015/68 supplementing Regulation (EU) No 167/2013 of the European Parliament and of the Council with regard to vehicle braking requirements for the approval of agricultural and forestry vehicles, October 2014.
3. Dimov I T, Sean McKee S. Monte Carlo Methods for Applied Scientists. World Scientific Press, 2004.
4. Forrer P. Brake systems in agricultural and forestry vehicles, http://www.paul-forrer.ch (accessed 07 May 2019).
5. Glišović J, Lukić J, Vanja Šušteršič V, Ćatić D. Development of tractors and trailers in accordance with the requirements of legal regulations. In: 9th International Quality Conference, Center for Quality, Faculty of Engineering, University of Kragujevac, June 2015, paper no. 3504: 193-201.
6. Gredeskul A B. O normativach effektivnosti tormoženija avtomobilej. Avtomobilnaja promyšlennost 1963; 6: 14-16, https://doi.org/10.1088/0031-9112/14/1/012.
7. Gredeskul A B, Fedosov V M, Skutnev V M. Opredelenie parametrov tormoznoj sistemy s regulatorom tormoznych sil. Avtomobilnaja promyšlennost 1975; 6: 24-26.
8. Haldex, Agricultural trailer product catalogue. Europe, Edition 1, 2015.
9. Kamiński Z. Distribution of braking forces in two-axle agricultural trailers. Teka Kom. Mot. Energ. Roln. 2005; 5: 80-86.
10. Kaminski Z, Miatluk M. Brake systems of road vehicles. Calculations. Bialystok: Wydawnictwo Politechniki Bialostockiej, 2005.
11. Kamiński Z. Simulation and experimental testing of the pneumatic brake systems of agricultural vehicles. Bialystok: Oficyna Wydawnicza Politechniki Białostockiej, 2012.
12. Kamiński Z, Kulikowski K. Determination of the functional and service characteristics of the pneumatic system of an agricultural tractor with mechanical brakes using simulation methods. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17(3): 355-364, https://doi.org/10.17531/ein.2015.3.5.
13. Kamiński Z. Mathematical modelling of the trailer brake control valve for simulation of the air brake system of farm tractors equipped with hydraulically actuated brakes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(4): 637-643.
14. Keyser DE, Hogan K. Hydraulic brake systems and components for off-highway vehicles and equipment. National Fluid Power Association Technical Paper Series 1992; I 92-1.4: 1-9.
15. Keyser DE. Full power hydraulic brake actuation, circuit design considerations for off-highway vehicles and equipment. In: 10th International Conference on Fluid Power - the Future for Hydraulics, Brugge, Belgium, 5-7 April 1993, edited by N. Way. Mechanical Engineering Publications, London.
16. Khaled M, Mahmoud R. Theoretical and experimental investigations one new adaptive duo servo drum brake with high and constant brake shoe factor, university Paderborn, 2005.
17. Knorr-Bremse, Agricultural and forestry vehicles. Brake equipment catalogue, Y206317 - (EN - Rev. 001), 2015.
18. Kroese DP, Taimre T, Botev ZI. Handbook of Monte Carlo Methods. New York, 2011. https://doi.org/10.1002/9781118014967
19. Lin M, Zhang W. Dynamic simulation and experiment of a full power hydraulic braking system. Journal of University of Science and Technology Beijing 2007; 29(10): 70-75.
20. Morton DP, Popova E. Monte Carlo simulations for stochastic optimization: Encyclopedia of Optimization. In: Floudas CA, Pardalos PM (eds) Monte Carlo simulations for stochastic optimization. Kluwer Academic Publishers, 2001; 1529-1537, https://doi.org/10.1007/0-306-48332-7_305.
21. Radlinski RW, Flick MA. Tractor and trailer brake system compatibility. SAE Transactions; paper no. 861942, 1986, https://doi.org/10.4271/861942.
22. Safim. Trailer brake valve, http://www.italgidravlika.ru/pdf_files/Safim/safim_11.pdf (accessed 15 May 2018).
23. Tang G, Zhao H, Wu J, Zhang Y. Optimization of Braking Force Distribution for Three-Axle Truck. SAE Technical Paper 2013-01-0414, 2013, https://doi.org/10.4271/2013-01-0414.
24. UN Economic Commission for Europe, ECE Regulation No. 13. Uniform provisions concerning the approval of vehicles of categories M, N and O with regard to braking, Geneva, Switzerland, 2001.
25. Venkataraman P. Applied Optimization with MATLAB Programming Wiley-Interscience. New York, 2001.
26. Wabco, FPB - Full Hydraulic Power Brake, Version 2/09, 2013.
27. Wabco, Off-highway. Overview technologies and products, Edition 2, Version 3, December 2016.
28. Wabco, Air braking system. Agriculture and forestry vehicles, Edition 11, Version 1, October 2017.

Zbigniew KAMIŃSKI
Paweł RADZAJEWSKI
Bialystok University of Technology
Faculty of Mechanical Engineering
ul. Wiejska 45C, 15-351 Białystok, Poland

Emails: radzajewskilech@wp.pl z.kaminski@pb.edu.pl

Article citation info:
QUEZADA DEL VILLAR AV, RODRÍGUEZ-PICÓN LA, PÉREZ-OLGUÍN IJC, MÉNDEZ-GONZÁLEZ LC. Stochastic modelling of the temperature increase in metal stampings with multiple stress variables and random effects for reliability assessment. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 654–661, http://dx.doi.org/10.17531/ein.2019.4.15.

Andrea Viridiana QUEZADA del VILLAR
Luis Alberto RODRÍGUEZ-PICÓN
Iván JC PÉREZ-OLGUÍN
Luis Carlos MÉNDEZ-GONZÁLEZ

STOCHASTIC MODELLING OF THE TEMPERATURE INCREASE IN METAL STAMPINGS WITH MULTIPLE STRESS VARIABLES AND RANDOM EFFECTS FOR RELIABILITY ASSESSMENT

OCENA NIEZAWODNOŚCI Z WYKORZYSTANIEM STOCHASTYCZNEGO MODELU WZROSTU TEMPERATURY W METALOWYCH WYTŁOCZKACH, UWZGLĘDNIAJĄCEGO WIELORAKIE ZMIENNE NAPRĘŻENIOWE ORAZ EFEKTY LOSOWE

Many products wear out over time even before they fail or stop working, therefore, through accelerated degradation tests one is able to make inferences about statistical parameters or the distributions of a product useful life. Since many devices experience different types of variation due to unobservable factors during the manufacturing processes or under certain operating conditions; these situations lead to the need in developing accelerated degradation models with several variables of acceleration and random effects. The proposed model in this paper, is a model based on the gamma process with random effects to have a better analysis of degradation. This model is applied to the analysis of the temperature increase of metal stampings that are affected by multiple explanatory variables. In addition, a statistical inference method based on a Bayesian approach is used to estimate the unknown parameters to then perform a reliability analysis after obtaining the first-passage time distributions.

***Keywords**: accelerated degradation test, gamma process, random effects, reliability assessment.*

Wiele produktów zużywa się z upływem czasu zanim nawet ulegną uszkodzeniu lub przestaną działać. Badania przyspieszonego starzenia pozwalają wyciągać wnioski na temat parametrów statystycznych lub rozkładów okresu użytkowania produktu. Wiele urządzeń podlega różnym rodzajom zmienności pod wpływem działania nieobserwowalnych czynników występujących podczas procesu produkcyjnego lub w pewnych warunkach pracy; sytuacje te wymagają opracowania modeli przyspieszonego starzenia uwzględniających wielorakie zmienne przyspieszenia oraz efekty losowe. Zaproponowany w przedstawionym artykule model opiera się na procesie gamma z efektami losowymi, dzięki czemu pozwala na lepszą analizę degradacji. Model ten zastosowano do analizy wzrostu temperatury w metalowych wytłoczkach, na które oddziałuje wiele zmiennych objaśniających. Ponadto do oszacowania nieznanych parametrów wykorzystano metodę wnioskowania statystycznego opartą na podejściu bayesowskim. Umożliwiło to analizę niezawodności po uzyskaniu rozkładów czasu pierwszego przejścia.

***Słowa kluczowe**: badanie przyspieszonego starzenia, proces gamma, efekty losowe, ocena niezawodności.*

1. Introduction

Currently, the level of quality and reliability of the products offered to customers is essential to maintain competitiveness in the market. Furthermore, to obtain highly reliable products, very large experimentation times are required, an alternative that mitigates this drawback are the accelerated degradation tests (ADT), in which a group of products suffer high levels of environmental stress [17]. In this way, it is possible to obtain the degradation measurements of a particular performance characteristic and the failure times in the shortest time possible. In addition, by including random effects on the variables [32], it is possible to describe the variation of the initial conditions of the devices as a function of the parameters of the model and to improve the accuracy of the predictions of reliability under conditions of normal use [22].

According to Peng & Tseng [26] in this situation the degradation models that are described by stochastic processes (gamma process, Wiener process, Gaussian inverse process), are useful for the analysis of the performance of one or several magnitudes that vary randomly as a function of time. Where some of its parameters can be affected and modified in order to incorporate the variability and the explanatory variables that affect the increase in degradation. Finally, with the correct specification of the model and the adequate supervision of the degradation processes, it is possible to predict the remaining useful life of the product, precise when maintenance or replacements are necessary or appropriate, be certain of the products useful life time, and provide greater reliability to the client [46,43].

Therefore, manufacturers must constantly innovate and produce value-added components [18] and in order to avoid or reduce failures, stochastic process models are naturally applicable and allow to counteract the previous problems. Among them is the Wiener process, [33] this stochastic process has been used to model the light intensity of LED lamps; Joseph & Yu [15] to improve reliability; Wang, et al. [42] proposed an adaptive method and a numerical example about

cracks (low intensity) due to fatigue; Barker & Newby [3] to describe the degradation of a multi-component system and develop an optimal non-periodic inspection policy; in areas such as medicine, to describe a series of biomarkers that represent the deterioration of systems over time, in a population of individuals infected with HIV to predict the residual time since the entry into the study until the moment in which a critical limit is reached [9]. Another model is the inverse Gaussian process which has been used by Wang & Xu [40] to fit the laser GaAs data; Peng [25] and Ye et al. [45] discussed when the underlying degradation follows the Gaussian inverse process; Zhang, et al. [50] describe a model to characterize the growth of depth defects by corrosion in underground energy pipelines; in fatigue cracks of structural components from metallic aircrafts, mainly in aluminum parts [19]; in the location of wireless sensor networks [28]; for the analysis of the degradation of industrial bearings [13]; in a series of biomarkers that represent the deterioration of the systems over time, in a population of individuals infected with HIV to predict the latency time from the moment of infection until the moment in which it is detected [9]. The Poisson process has also been presented as a model to deal with the reliability estimation of systems and processes. Andrzejcza et al. [2] presented a Poisson based model to estimate the cost of corrective maintenance of public mass transport vehicles.

On the other hand, the gamma process has been used extensively in the literature, this due to its important characteristics that it can be used when a large number of product failures is caused by the impact of external random factors that tend to be very small, with independent increases, with base of zero and up to infinity (not negative) and where performance can only decrease with respect to time [43, 16 10]. Therefore, it is adequate to model gradual damage where a stochastic monotonous deterioration accumulates [35], the increments are stationary [37]. Some applications have been presented as worn-out, with propagations of fissures, crack growth, erosion, consumption, creep, swelling, degrading health index, corrosion, consumption and fatigue, among other factors [29,36]; by Iervolino, et al. [14] to model the effect produced by earthquakes; van Noortwijk [36] studied the application in maintenance Wang, et al. [41] proposed an adaptive method applied to the growth of fatigue cracks; Bordes, et al. [4] considered a degradation model that consists of two independent processes, in addition they illustrated their method through a study and an application to a set of real data presented in the article by Takeda & Suzuki [31]; Pan, et al. [21] proposed reliability models for systems with two degrading components depending on an example of a railway.

This article presents a case study related to the increase of temperature in metal stampings which were exposed to different levels of electric current, apart from considering different explanatory variables during an ADT. The analysis of this case study is carried using the gamma process by considering the use of a life-stress relationship such as the exponential link function. This life-stress relationship is commonly used when the case of multiple stress variables is presented, some important applications can be found in: Park & Padgett [24] that suggested a hyper-cuboidal volume approach as a measure of acceleration that can incorporate several acceleration variables. A special case that includes the Weibull model and the law of power [20]. By considering the proportional hazard model [6] as an extended model with a weak link, with hyper-cuboidal volume and the consideration of random effects. In different applications such as bridge beam data that includes degradation due to the entry of chloride ions [39]. To compare a set of fatigue-cracks growth data [47]; to evaluate the effectiveness of laser photocoagulation to delay visual loss in patients with diabetic retinopathy [49]. As a model that has applications in maintenance [7], and in the degradation of light intensity of an electronic device [48]. On the other hand, random effects are considered in the gamma process model in order to include the unit to unit variation. The parameters of interest are estimated based on a Bayesian approach via Markov chain Monte Carlo (MCMC). Likewise, a comparison was made with other stochastic processes in order to obtain the one that best fits the data that was used.

The rest of this document is organized as follows. In Section 2, the characteristics, functions and the first-passage time distribution of the gamma process are presented. First, the gamma process is introduced with random effects, and then the exponential link relationship is presented as a life stress function that best fits the explanatory variables that affect the case study. In Section 3, an estimation scheme for the gamma process parameters is presented, under a Bayesian analysis with MCMC approach. In Section 4, a case study based on the temperature increase of metal stampings is provided to illustrate the application and utility of the proposed model; the estimation of the model was made based on Bayesian inference, to obtain the first passage time distributions; finally, a comparison with other stochastic processes is discussed. The conclusions can be seen in Section 5.

2. Stochastic model based on the gamma process

2.1. General Characteristics

In this paper, it is considered that the degradation of a performance characteristic at time t $\{Z(t);t \geq 0\}$ is governed by a gamma process. The gamma process has been widely used in the literature as described in Section 1. The main properties of this process are:

1. $Z(0) \equiv 0$;
2. $Z(t)$ has independent increments.
3. For any $t > s; Z(t) - Z(s) = \Delta Z(t)$ follows a gamma distribution $Ga(v(t-s),u) = Ga(v\Delta t,u)$

where $v(t);t \geq 0$ is a non-negative increasing function with $t \geq 0$ and $v(0) \equiv 0$, also known as the shape function. And, $u > 0$ is the scale parameter. Supposing that the stochastic gamma process describe the degradation level of some performance characteristic at time t, then the probability density function (PDF) can be denoted as in (1), with mean $vt \cdot u$, variance $vt \cdot u^2$, and compact notation as $Z(t) \sim Ga(vt,u)$.

$$f(Z(t)|vt,u) = \frac{Z(t)^{vt-1}}{u^{vt}\Gamma(vt)} e^{-\left(\frac{Z(t)}{u}\right)}. \tag{1}$$

The moment of failure is an important characteristic that can be obtained from (1) and occurs when the degradation $Z(t)$ reaches a critical level of degradation ω [5]. This moment of failure is a random variable T_ω with expression denoted as:

$$T_\omega = inf\{Z(t) \geq \omega\}. \tag{2}$$

On the other hand, the cumulative distribution function (CDF) of T_ω can be derived based on:

$$\begin{aligned} F_{Ga}(t_\omega) &= P(T_\omega \leq t_\omega) = P(Z(t) \geq \omega) \\ &= \int_\omega^\infty f_{Z(t)}(z)dz = \int_\omega^\infty \frac{z^{vt-1}}{u^{vt}} e^{-\left(\frac{z}{u}\right)} dz \\ &= \frac{1}{\Gamma(vt)} \int_{\omega/u}^\infty x^{vt-1} e^{-x} dx, \end{aligned}$$

where $x = z / u$, the integral in the last previous equation can be simplified by considering the incomplete gamma function, $\Gamma(a,b) = \int_b^\infty \zeta^{a-1} e^{-\zeta} d\zeta$. Thus, under the notation $a = vt$ and $b = \omega / u$, the CDF results as denoted in (3).

$$F_{Ga}(t_\omega) = \frac{\Gamma(vt, \omega / u)}{\Gamma(vt)}. \tag{3}$$

Now suppose that an ADT has been performed to N devices that are observed during M inspections until the test termination time T. Then, $Z_i(t_j)$ degradation measurements are observed for the trajectories $i = 1,2,\ldots,N$ at the corresponding times $t_j, j = 1,2,\ldots,M$. Considering the independent increment property of the gamma process, and $\Delta Z_i(t_j) = Z_i(t_j) - Z_i(t_{j-1}), t_0 = 0$ and $\Delta t_j = t_j - t_{j-1}$, thus the variable $\Delta Z_i(t_j)$ is governed by a gamma process model as denoted in (4) with compact notation $\Delta Z_i(t_j) \sim Ga(v\Delta t_j, u)$.

$$f\left(\Delta Z_i(t_j) | v\Delta t_j, u\right) = \frac{\Delta Z_i(t_j)^{v\Delta t_j} e^{-\left(\frac{\Delta Z_i(t_j)}{u}\right)}}{u^{v\Delta t_j} \Gamma(v\Delta t_j)}. \tag{4}$$

2.2. Random effects in the gamma process

Random effects are considered in the gamma process when it is suspected that non observable factors can cause variations in the observed degradation. Specifically random effects are incorporated to describe unit to unit variability, *i.e.,* heterogeneity among unities under test [46,10]. According to Rodríguez Picón, et al. [30], when performing a degradation test each unit under test may be affected by different sources of variation which denotes the need of incorporate random effects in the model. In the case of the gamma process, random effects can be incorporated as a function of the scale parameter. Then, it is considered that u_i is a random variable for the $i = 1,2,\ldots,N$ devices under test. It can be noted that as the mean of the gamma process is described as $vt \cdot u$ and the variance is described as $vt \cdot u^2$, the randomness of u_i have an impact over the mean degradation and the variance degradation. This means that it is expected that the degradation paths exhibit a large variation among paths (the mean degradation is a function of u_i) and a large variation within each path (the variance is also a function of u_i).

In the literature, it has been found that the random scale parameter (u_i) is described by a gamma distribution $Ga(\delta, \varphi)$ [38]. This model is known as the classical gamma process model with random effects. By considering this, the PDF of the degradation $\Delta Z_i(t_j)$ is represented as:

$$f\left(\Delta Z_i(t_j)\right) = \int_0^\infty f_{Ga}\left(\Delta Z_i(t_j) | v\Delta t_j, u_i\right) \cdot f_{Ga}(u_i | \delta, \varphi) du_i. \tag{5}$$

Thus, the CDF of the lifetime when the degradation path reaches the critical level of degradation ω can be obtained by solving the integral presented in (6):

$$F_R(t_\omega) = \int_0^\infty F_{Ga}(t_\omega) \cdot f_{Ga}(u_i | \delta, \varphi) du_i. \tag{6}$$

2.3. The exponential link relation as a function of the shape parameter

The ADTs are used to accelerate the degradation process and thus accelerate the failure time in the aims of obtaining a reliability estimation in less time. Of course, these type of tests consist in submitting a device to high levels of stress during a determined period of time, which allows to obtain the degradation increments that lead to the failure in a short time [34]. Furthermore, a product may be exposed to multiple covariates (s) that affects the degradation process, such as temperature, humidity, voltage, etc., and other characteristics such as material type, geometrical characteristics, etc. Then, it is important to incorporate such covariates in the model. Specifically as a function of a parameter of the proposed model [46].

The exponential link function has been used in the cases when multiple covariates have an effect on the performance of a characteristic of interest. This function relates a life characteristic and the stress variables through exponential functions [11]; the general form of this function is presented in (7) [44]:

$$h(s_k) = \beta_0 e^{\beta_k s_k}, \tag{7}$$

where $h(s_k)$ can repesent a life characteristic that is observed under the effect of $s_k = s_1, s_2, \ldots s_p$ stress variables, where p represents the number of covariates. On the other hand, β_0 and $\beta_k (k = 1,2,\ldots p)$ are constant parameters to be estimated [23]. In the case of the gamma process, in the literature it has been found that the shape parameter describes the effect of stress on the performance of products, such that v can relate the effect of $h(s_k)$. Thus, it is considered that $v(s_k) = h(s_k) = \beta_0 e^{\beta_k s_k}$ for $k = 1,2,\ldots p$. In this way, the PDF of the gamma process with the kth stress variable can be expressed as in (8):

$$f\left(\Delta Z_i(t_j) | \delta, \varphi, \beta_0, \beta_k\right) = \int_0^\infty f_{Ga}\left(\Delta Z_i(t_j) | v(s_k)\Delta t_j, u_i\right) \cdot f_{Ga}(u_i | \delta, \varphi) du_i;$$
$$k = 1,2,\ldots p;\ i = 1,2,\ldots,N;\ j = 1,2,\ldots,M \tag{8}$$

The CDF of the moment of failure when the degradation path reaches the critical value ω can be found by solving the integral in (9):

$$F(t_\omega | s_k) = \int_0^\infty F_{Ga}\left(t_\omega | v(s_k) t_j, u_i\right) \cdot f_{Ga}(u_i | \delta, \varphi) du_i \tag{9}$$

According to Lawless & Crowder[16], the integral in (9) results in terms of the Fisher distribution. Then, by considering the exponential link function, the CDF of the first passage times result as in (10):

$$F(t_\omega | s_k) = 1 - F_{v(s_k)t_j, 2\delta}\left(\frac{\delta\omega}{\varphi v(s_k) t_j}\right) \tag{10}$$

3. Estimation of parameters

Certainly, it results of interest to estimate the parameters $(\beta_0, \beta_k, \delta, \varphi)$ from the model presented in (8), such that it is possible to obtain reliability estimations through the function presented in (10). However, it can be noted from the integral presented in (8) that the classical methods of estimation, such as the maximum likelihood estimation (MLE) method, result too complicated to implement given the complexity of the function in (8). On the other hand, in the last

years there have been important advances in numeric approximation techniques to deal with complex functions via MCMC [12]. Specifically, the estimation of complex functions via a Bayesian approach with MCMC has been presented as an important alternative. Definitely, the implementation of a numerical technique such as a Bayesian MCMC estimation scheme to solve complex functions requires the use of a specialized software. Fortunately, there several open source softwares such as OpenBUGS and R which can be used. In this paper, we consider an estimation scheme based on a Bayesian approach with MCMC to estimate the parameters in the model (8). This approach has the practical advantage that it can incorporate subjective information in a natural way when there is little information about the historical behavior of the parameters on interest.

In Figure 1, we present the general estimation scheme based on a Bayesian approach. For the parameters of interest $(\beta_0, \beta_k, \delta, \varphi)$, we consider non-informative prior distributions (which are denoted as $\pi(\beta_0), \pi(\beta_k), \ldots$), given that no prior knowledge of these parameters is available. Specifically, the next prior distributions were considered:

$$\pi(\beta_k) \sim N(0, \tau^2); k = 1,2,\ldots p$$

$$\pi(\beta_0) \sim N(0, \tau^2)$$

$$u_i \sim Ga(\delta, \varphi)$$

$$\pi(\delta); \delta \sim f_{Ga}(0.01, 0.01)$$

$$\pi(\varphi); \varphi \sim f_{Ga}(0.01, 0.01)$$

Non-informative prior normal distributions $N(0, \tau^2)$ were considered for the parameters of the exponential link function $(\beta_0, \beta_k; k = 1,2,\ldots p)$ with mean 0 and $\tau^2 = 0.001$, where $\tau^2 = 1/\sigma^2$ is know as a precision parameter. While, non-informative prior gamma distributions were considered for the parameters of the random scale $u_i \sim Ga(\delta, \varphi)$, with shape parameters $a_\delta = 0.01, a_\varphi = 0.01$ and scale parameters $b_\delta = 0.01, b_\varphi = 0.01$, respectively. Considering these prior distributions and the likelihood function of (8), the posterior distribution can be expressed as in (11). This estimation scheme was programmed in OpenBUGS and solved via MCMC.

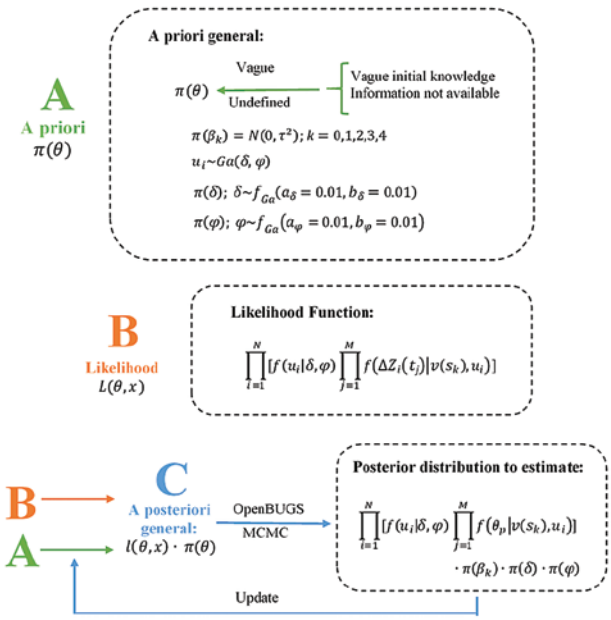


Fig. 1. Bayesian estimation scheme for the stochastic gamma process with random effects and link exponential function. Based on [27]

$$f(\beta_0, \beta_k, \delta, \varphi | \Delta Z_i(t_j))$$
$$\propto \prod_{i=1}^{N}[f(u_i|\delta,\varphi)\prod_{j=1}^{M} f(\Delta Z_i(t_j)|v(s_k), u_i)] \cdot \pi(\beta_0) \cdot \pi(\beta_k) \cdot \pi(\delta) \cdot \pi(\varphi);$$
$$k = 1,2,\ldots p; i = 1,2,\ldots,N; j = 1,2,\ldots,M \tag{11}$$

4. Case study

The case study consists in an ADT performed to obtain the temperature increase of metal stampings that are incorporated in printed circuit boards (PCB). These PCBs are used in the fuse box of a certain automobile. The ADT was performed by applying different levels of current into the stamping as 60, 80 and 100 amperes. In addition, three types of materials were used in the test, which are: CDA151, CDA210 and CDA425, and four lengths of the stamping were also considered as 100, 150, 200 and 250 mm. This test configuration resulted in a total of five replicates each with 28 combinations of the described levels of the three factors. A total of 140 metal stampings, *i.e.* $i = 1,2,\ldots,140$, were subjected to the different levels of current during 30 minutes. Thus, all the devices were observed simultaneously at $t_0 = 30, t_1 = 60, t_2 = 90, t_3 = 120, t_4 = 150$ and $t_5 = 180$. Given that the fuse box has a limit of temperature of 170°C, and for safety reasons,

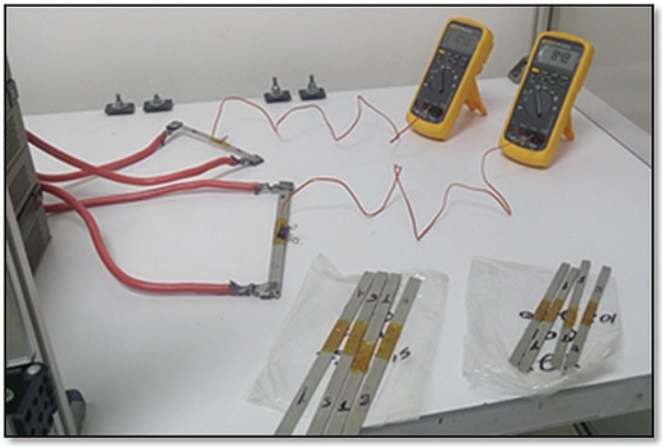


Fig. 2. Test configuration to obtain the temperature increase of the metal stampings

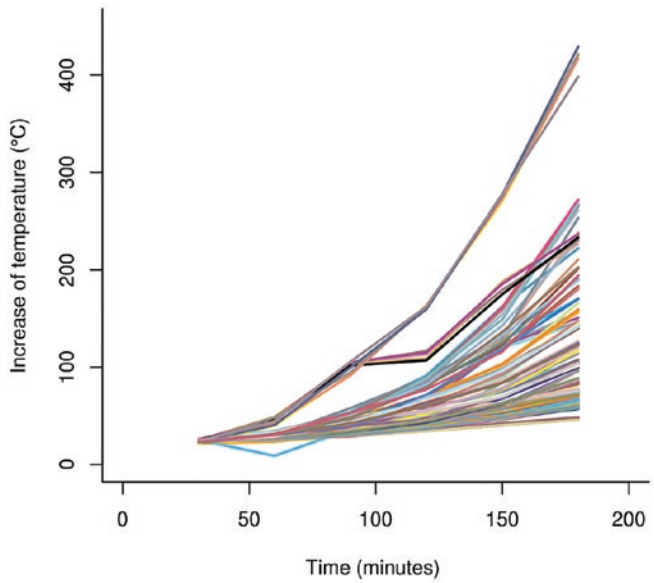


Fig. 3. Temperature increase paths for a sample of metal stampings

the maximum allowance for the temperature was set at 145°C (maximum design temperature). In Figure 2, the general configuration of the test is presented, the material type and the test equipment can also be seen. Furthermore, as can be noted in Figure 2 all the temperature readings were obtained in the middle of the stamping, this in the aims of obtaining homogenous readings for all the stampings under test. In Figure 3, a sample of paths of the temperature increase of different metal stampings is presented.

4.2. Estimation of the model

Considering the stochastic gamma process with random effects and the exponential link function for the current, material type and the length of the stamping, then the posterior distribution described in (11) results in:

$$f\left(\beta_0,\beta_k,\delta,\varphi|\Delta Z_i\left(t_j\right)\right)$$
$$\propto \prod_{i=1}^{N}[f\left(u_i|\delta,\varphi\right)\prod_{j=1}^{M}f\left(\Delta Z_i\left(t_j\right)|v\left(s_k\right),u_i\right)]\cdot\pi\left(\beta_0\right)\cdot\pi\left(\beta_k\right)\cdot\pi\left(\delta\right)\cdot\pi\left(\varphi\right);$$
$$k=1,2,3;\, i=1,2,\ldots,140;\, j=1,2,3,4,5$$

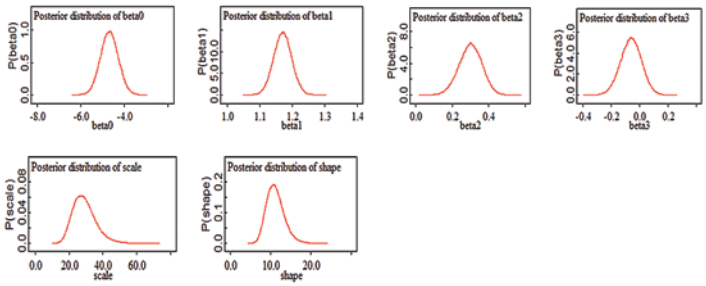


Fig. 4. Posterior distributions for $\left(\hat{\beta}_0,\hat{\beta}_1,\hat{\beta}_2,\hat{\beta}_3,\hat{\delta}\text{ and }\hat{\varphi}\right)$

Table 1. Obtained estimations of parameters under the Bayesian approach

	Mean	Sd	MC error	$t_{0.025}$	Median	$t_{0.975}$
$\hat{\beta}_0$	-4.688	0.4074	0.004095	-5.479	-4.69	-3.887
$\hat{\beta}_1$	1.171	0.02764	2.99E-04	1.117	1.171	1.225
$\hat{\beta}_2$	0.3002	0.06246	9.42E-04	0.1745	0.3014	0.4188
$\hat{\beta}_3$	-0.06096	0.07387	8.00E-04	-0.2082	-0.06007	0.08192
$\hat{\delta}$	28.99	6.769	0.1169	17.94	28.29	44.48
$\hat{\varphi}$	11.1	2.185	0.0356	7.403	10.91	15.96

The estimation scheme proposed in Figure 1 was considered for the estimation of the parameters of the model described above $\left(\beta_0,\beta_1,\beta_2,\beta_3,\delta,\varphi\right)$. According to the proposed scheme, the OpenBUGS software was used for the implementation of the MCMC method. The algorithm was constructed considering the model presented in (8) and non-informative prior distributions as follows: non-informative normal distributions were considered for the parameters of the link exponential function as $\beta_0,\beta_1,\beta_2,\beta_3 \sim N\left(0,0.0001\right)$. In addition, non-informative prior gamma distributions were considered for the parameters of the random effects parameter u_i as $\delta,\varphi \sim Ga\left(0.01,0.01\right)$. A total of 400,000 iterations were considered, from which a total of 50,000 were disregarded for burn-in purposes. In Table 1, a summary of the obtained estimations is presented. While, in Figure 4 the posterior distributions for $\left(\beta_0,\beta_1,\beta_2,\beta_3,\delta,\varphi\right)$ are presented.

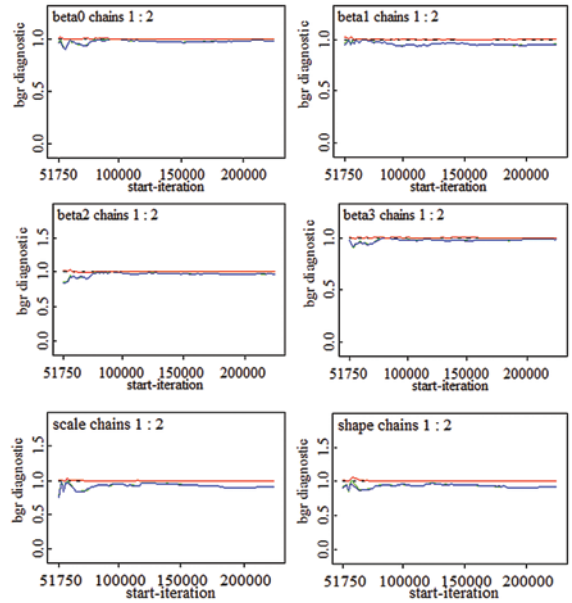


Fig. 5. BGR diagnostic for the parameters of interest

The Brooks-Gelman Rubin (BGR) diagnostic statistic was used in the aims of determining convergence in the estimation of parameters. For this, two chains of initial values were defined for the parameters of interest. The first chain of initial values was defined as $\beta_0=0.005,\beta_1=0.006,\beta_2=0.007,\beta_3=0.002,\delta=2$ and $\varphi=0.001$, while the second chain the next initial values were considered $\beta_0=0.1,\beta_1=0.1,\beta_2=0.1,\beta_3=0.1,\delta=1$ and $\varphi=0.1$. In Figure 5, the BGR graphs are presented for the parameters of interest under the two chains of initial values. Given that for all the

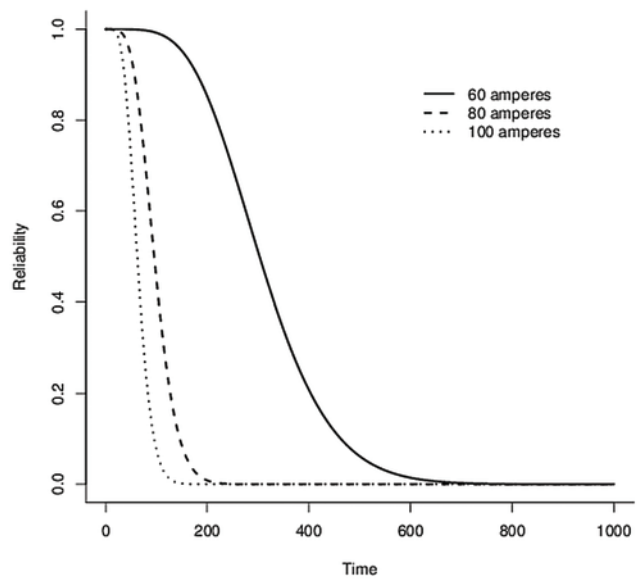


Fig. 6. Reliability functions under different levels of current

parameters of interest, the behavior of the two chains is around 1, then it can be said that convergence was achieved.

4.3. First passage time distribution

Considering the estimated parameters of interest that are presented in table 1, it results necessary to characterize the first passage time distributions when the degradation trajectories cross the critical level 145°C. As presented in equation (10) the CDF of the first passage times is described as:

$$F\left(t_{\omega}|s_k\right)=1-F_{v(s_k)t_j,2\delta}\left(\frac{\delta\omega}{\varphi v\left(s_k\right)t_j}\right)$$

Thus, the reliability function can be described as:

$$R\left(t_{\omega}|s_k\right)=F_{v(s_k)t_j,2\delta}\left(\frac{\delta\omega}{\varphi v\left(s_k\right)t_j}\right) \tag{12}$$

where the parameter $v\left(s_k\right), k=1,2,3$ is related to the link exponential function as $v\left(s_k\right)=\beta_0 e^{1.171s_1+0.3002s_2-0.06096s_3}$, where $s_1=current$, $s_2=material\,type$ and $s_3=length$. For this case study, we coded the material types as $CDA151=1$, $CDA210=2$ and $CDA425=3$. Of course different first passage time reliability functions can be characterized depending of the levels of $s_k, k=1,2,3$. In this case, we are interested in the material type $CDA151$ with coded value as 1, and the stamping length of 100. Under these two fixed levels of these two factors and the different levels of the current, we evaluated the function presented in equation (12). In figure 6, the different reliability functions are presented.

The effect of the current over the temperature of the stamping can be noted in the behavior of the reliability functions. In addition, the mean time to failure (MTTF) can be obtained as:

$$MTTF=\int_0^{\infty}R\left(t_{\omega}|s_k\right).$$

By considering the reliability function in (12) and the previously discussed levels of the material type and stamping length, the MTTF under the different levels of current were obtained as $MTTF=313.8065\,minutes$ for 60 amperes, $MTTF=98.74441\,minutes$ for 80 amperes and $MTTF=64.33426\,minutes$ for 100 amperes.

4.4. Comparison with other stochastic process

Besides the gamma process, we also considered the Wiener and the inverse Gaussian process to model the dataset obtained from the ADT. Equivalent stochastic degradation models with random effects from these two stochastic processes were considered, specifically the models proposed by Cheng & Peng [5] and Ye & Chen [46]. The comparison was carried out based on the Akaike information criterion (AIC), which is obtained via the formula $\mathrm{AIC}=(-2)\mathrm{loglikelihood}+2\,(\text{number of parameters})$ [8,1].

The obtained AIC values for the three stochastic processes resulted in: for the Wiener process a value of 239.5724, for the gamma process a value of 232.2521 was obtained, while for the inverse Gaussian process a value of 234.3279 was obtained. Based on these results, it can be noted that the stochastic gamma process is more adequate to model the degradation dataset of the presented case study. This, given that the AIC value was the lowest compared to the other two stochastic processes.

5. Conclusions

In this paper, a stochastic degradation model based on the gamma process was proposed, in which random effects and multiple stress variables are incorporated. The estimation of parameters was carried out using a Bayesian approach considering MCMC. The proposed model as well as the estimation scheme were implemented in a case study that consisted in the increase of temperature on metal stamping under different electric current levels, different types of material and different plate lengths were also considered. Given that non-informative a priori distributions were considered for the parameters of interest, two chains of initial values were defined in order to evaluate the convergence using a BGR graph. Given the behavior of the BGR graph it was noted that convergence was observed in all parameters. With the parameters obtained it was possible to estimate the reliability under the different electric current levels for the material type CDA-151 and a length of 100 mm, in the same way it was possible to obtain the MTTF values. The effect of the current and the increase in the temperature on the metal stamping can be seen both in the reliability functions and in the MTTF's. Finally, a comparison was made with the Wiener and inverse Gaussian processes based on the AIC value, given the obtained results it can be said that the gamma process turned out to be the best. As a future work, other characteristics of the stamping may be considered, to add them to the exponential link function, such as the width and thickness. These characteristics can be easily added to the proposed model.

References

1. Akaike H. Factor analysis and AIC. In: Parzen E, Tanabe K, Kitagawa G, editors. Selected Papers of Hirotugu Akaike. New York, NY: Springer New York 1998; 371-386, https://doi.org/10.1007/978-1-4612-1694-0_29.
2. Andrzejczak K, Młyńczak M, Selech J. Poisson-distributed failures in the predicting of the cost of corrective maintenance. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 602-609, https://doi.org/10.17531/ein.2018.4.11.
3. Barker CT, Newby MJ. Optimal non-periodic inspection for a multivariate degradation model. Reliability Engineering & System Safety 2009; 94(1): 33-43, https://doi.org/10.1016/j.ress.2007.03.015.
4. Bordes L, Paroissin C, Salami A. Parametric inference in a perturbed gamma degradation process. Communications in Statistics - Theory and Methods 2016; 45(9): 2730-2747, https://doi.org/10.1080/03610926.2014.892133.
5. Cheng Y-S, Peng C-Y. Integrated Degradation Models in R Using iDEMO. Journal of Statistical Software 2012; 49(2): 22, https://doi.org/10.18637/jss.v049.i02.
6. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society: Series B (Methodological) 1972; 34(2): 187-202, https://doi.org/10.1111/j.2517-6161.1972.tb00899.x.
7. Crowder M, Lawless J. On a scheme for predictive maintenance. European Journal of Operational Research 2007; 176(3): 1713-1722, https://doi.org/10.1016/j.ejor.2005.10.051.

8. deLeeuw J. Introduction to Akaike (1973) Information Theory and an Extension of the Maximum Likelihood Principle. In: Kotz S, Johnson NL, editors. Breakthroughs in Statistics: Foundations and Basic Theory. New York, NY: Springer New York 1992; 599-609, https://doi.org/10.1007/978-1-4612-0919-5_37.
9. Doksum KA, Normand S-LT. Gaussian models for degradation processes-part I: Methods for the analysis of biomarker data. Lifetime Data Analysis 1995; 1(2): 131-144, https://doi.org/10.1007/BF00985763.
10. Esary JD, Marshall AW. Shock models and wear processes. The Annals of Probability 1973; 1(4): 627-649, https://doi.org/10.1214/aop/1176996891.
11. Fan T-H, Wang W-L, Balakrishnan N. Exponential progressive step-stress life-testing with link function based on Box-Cox transformation. Journal of Statistical Planning and Inference 2008; 138(8): 2340-2354, https://doi.org/10.1016/j.jspi.2007.10.002.
12. Gelfand AE, Smith AFM. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association 1990; 85(410): 398-409, https://doi.org/10.1080/01621459.1990.10476213.
13. Hong S, Zhou Z. Application of gaussian process regression for bearing degradation assessment. 2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining (ISSDM2012) 2012; 644-648.
14. Iervolino I, Giorgio M, Chioccarelli E. Gamma degradation models for earthquake-resistant structures. Structural Safety 2013; 45: 48-58, https://doi.org/10.1016/j.strusafe.2013.09.001.
15. Joseph VR, Yu IT. Reliability improvement experiments with degradation data. IEEE Transactions on Reliability 2006; 55(1): 149-157, https://doi.org/10.1109/TR.2005.858096.
16. Lawless J, Crowder M. Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Analysis 2004; 10(3): 213-227, https://doi.org/10.1023/B:LIDA.0000036389.14073.dd.
17. Meeker WQ, Escobar LA. Statistical Methods for Reliability Data. New York: JOHN WILEY & SONS, INC 1998.
18. Miranda AV. La industria automotriz en México: Antecedentes, situación actual y perspectivas. Contaduría y administración 2007; (221): 1-38.
19. Mohanty S, Chattopadhyay A, Peralta P, Das S, Willhauck C. Fatigue Life Prediction Using Multivariate Gaussian Process. 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 16th AIAA/ASME/AHS Adaptive Structures Conference,10th AIAA Non-Deterministic Approaches Conference, 9th AIAA Gossamer Spacecraft Forum, 4th AIAA Multidisciplinary Design Optimization Specialists Conference. Structures, Structural Dynamics, and Materials and Co-located Conferences: American Institute of Aeronautics and Astronautics 2008; 1-14, https://doi.org/10.2514/6.2008-1837.
20. Padgett WJ, Durham SD, Mason AM. Weibull Analysis of the Strength of Carbon Fibers Using Linear and Power Law Models for the Length Effect. Journal of Composite Materials 1995; 29(14): 1873-1884, https://doi.org/10.1177/002199839502901405.
21. Pan Z, Sun Q, Feng J. Reliability modeling of systems with two dependent degrading components based on gamma processes. Communications in Statistics - Theory and Methods 2016; 45(7): 1923-1938, https://doi.org/10.1080/03610926.2013.870201.
22. Pan ZF, Jing; Sun, Quan. Lifetime distribution and associated inference of systems with multiple degradation measurements based on gamma processes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (2): 307-313, https://doi.org/10.17531/ein.2016.2.20.
23. Park C, Padgett WJ. Stochastic degradation models with several accelerating variables. IEEE Transactions on Reliability 2006; 55(2): 379-390, https://doi.org/10.1109/TR.2006.874937.
24. Park C, Padgett WJ. Cumulative Damage Models for Failure with Several Accelerating Variables. Quality Technology & Quantitative Management 2007; 4(1): 17-34, https://doi.org/10.1080/16843703.2007.11673132.
25. Peng C-Y. Inverse Gaussian processes with random effects and explanatory variables for degradation data. Technometrics 2015; 57(1): 100-111, https://doi.org/10.1080/00401706.2013.879077.
26. Peng C, Tseng S. Mis-specification analysis of linear degradation models. IEEE Transactions on Reliability 2009; 58(3): 444-455, https://doi.org/10.1109/TR.2009.2026784.
27. Peng W, Li Y-F, Yang Y-J, Huang H-Z, Zuo MJ. Inverse Gaussian process models for degradation analysis: A Bayesian perspective. Reliability Engineering & System Safety 2014; 130: 175-189, https://doi.org/10.1016/j.ress.2014.06.005.
28. Richter P, Toledano-Ayala M. Revisiting Gaussian process regression modeling for localization in wireless sensor networks. Sensors 2015; 15(9): 22587-22615, https://doi.org/10.3390/s150922587.
29. Rodríguez Picón LA, Rodríguez Borbón MI, Flores Ortega A. Estimación de confiabilidad para productos con dos características de desempeño basada en un modelo de degradación estocástico bivariado. CULCyT 2015; (57): 64-75.
30. Rodríguez Picón LA, Rodríguez Picón AP, Méndez González LC, Rodríguez Borbón MI, Alvarado Iniesta A. Degradation modeling based on gamma process models with random effects. Communications in Statistics - Simulation and Computation 2018; 47(6): 1796-1810.
31. Takeda E, Suzuki N. An empirical model for device degradation due to hot-carrier injection. IEEE Electron Device Letters 1983; 4(4): 111-113, https://doi.org/10.1109/EDL.1983.25667.
32. Tseng S-T, Yu H-F. A termination rule for degradation experiments. IEEE Transactions on Reliability 1997; 46(1): 130-133, https://doi.org/10.1109/24.589938.
33. Tseng ST, Tang J, Ku IH. Determination of burn-in parameters and residual life for highly reliable products. Naval Research Logistics 2003; 50(1): 1-14, https://doi.org/10.1002/nav.10042.
34. Tseng ST, Tsai CC, Balakrishnan N. Optimal Sample Size Allocation for Accelerated Degradation Test Based on Wiener Process 2011: 1-18, https://doi.org/10.1002/0471667196.ess7151.
35. van Noortwijk J, Kok M, Cooke R. Optimal maintenance decisions for the sea-bed protection of the Eastern-Scheldt barrier. In: Cooke R, Mendel M, Vrijling H, editors. Engineering Probabilistic Design and Maintenance for Flood Protection. Springer, Boston, MA 1997: 25-56, https://doi.org/10.1007/978-1-4613-3397-5_2.
36. van Noortwijk JM. A survey of the application of gamma processes in maintenance. Reliability Engineering & System Safety 2009; 94(1): 2-21, https://doi.org/10.1016/j.ress.2007.03.019.
37. van Noortwijk JM, Cooke RM, Kok M. A Bayesian failure model based on isotropic deterioration. European Journal of Operational Research 1995; 82(2): 270-282, https://doi.org/10.1016/0377-2217(94)00263-C.
38. Wang H, Xu T, Mi Q. Lifetime prediction based on Gamma processes from accelerated degradation data. Chinese Journal of Aeronautics 2015; 28(1): 172-179.

https://doi.org/10.1016/j.cja.2014.12.015
39. Wang X. A pseudo-likelihood estimation method for nonhomogeneous gamma process model with random effects. Statistica Sinica 2008; 18(3): 1153-1163.
40. Wang X. Wiener processes with random effects for degradation data. Journal of Multivariate Analysis 2010; 101(2): 340-351, https://doi.org/10.1016/j.jmva.2008.12.007.
41. Wang X, Balakrishnan N, Guo B, Jiang P. Residual life estimation based on bivariate non-stationary gamma degradation process. Journal of Statistical Computation and Simulation 2015; 85(2): 405-421, https://doi.org/10.1080/00949655.2013.824448.
42. Wang X, Guo B, Cheng Z. Residual life estimation based on bivariate Wiener degradation process with time-scale transformations. Journal of Statistical Computation and Simulation 2014; 84(3): 545-563, https://doi.org/10.1080/00949655.2012.719026.
43. Wang X, Xu D. An inverse Gaussian process model for degradation data. Technometrics 2010; 52(2): 188-197, https://doi.org/10.1198/TECH.2009.08197.
44. Wolstenholme LC. A nonparametric test of the weakest-link principle. Technometrics 1995; 37(2): 169-175, https://doi.org/10.1080/00401706.1995.10484301.
45. Ye Z-S, Chen L-P, Tang LC, Xie M. Accelerated degradation test planning using the inverse gaussian process. IEEE Transactions on Reliability 2014; 63(3): 750-763, https://doi.org/10.1109/TR.2014.2315773.
46. Ye Z-S, Chen N. The inverse Gaussian process as a degradation model. Technometrics 2014; 56(3): 302-311, https://doi.org/10.1080/00401706.2013.830074.
47. Ye Z-S, Xie M, Tang L-C, Chen N. Semiparametric estimation of gamma processes for deteriorating products. Technometrics 2014; 56(4): 504-513, https://doi.org/10.1080/00401706.2013.869261.
48. Ye Z-S, Xie M, Tang L-C, Shen Y. Degradation-based burn-in planning under competing risks. Technometrics 2012; 54(2): 159-168, https://doi.org/10.1080/00401706.2012.676946.
49. Zeng D, Lin DY, Yin G. Maximum likelihood estimation for the proportional odds model with random effects. Journal of the American Statistical Association 2005; 100(470): 470-483, https://doi.org/10.1198/016214504000001420.
50. Zhang S, Zhou W, Qin H. Inverse Gaussian process-based corrosion growth model for energy pipelines considering the sizing error in inspection data. Corrosion Science 2013; 73: 309-320, https://doi.org/10.1016/j.corsci.2013.04.020.

Andrea Viridiana QUEZADA DEL VILLAR
Luis Alberto RODRÍGUEZ-PICÓN
Iván JC PÉREZ-OLGUÍN
Luis Carlos MÉNDEZ-GONZÁLEZ
Department Industrial Engineering and Manufacturing
Autonomous University of Ciudad Juarez
Av. Plutarco Elías Calles 1210, Fovissste Chamizal
32310 Ciudad Juárez, Chihuahua, México

E-mails: al171779@alumnos.uacj.mx, luis.picon@uacj.mx, ivan.perez@uacj.mx, luis.mendez@uacj.mx

Article citation info:
BORUCKA A, NIEWCZAS A, HASILOVA K. Forecasting the readiness of special vehicles using the semi-Markov model. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 662–669, http://dx.doi.org/10.17531/ein.2019.4.16.

Anna BORUCKA
Andrzej NIEWCZAS
Kamila HASILOVA

FORECASTING THE READINESS OF SPECIAL VEHICLES USING THE SEMI-MARKOV MODEL

PROGNOZOWANIE GOTOWOŚCI POJAZDÓW SPECJALNYCH NA PODSTAWIE MODELU SEMI-MARKOWA*

The vehicle exploitation system, consisting of statistically identical objects that perform intervention tasks, not subject to systematic changes, can be modelled as a stationary stochastic process. Such a model allows to determine the probabilistic indicators of current and boundary readiness of the system. This article presents the use of the semi-Markov process, based on three operating states: operation, ready-to-be-used and repair, to study a transport system consisting of special vehicles. On the example of a sample consisting of police patrol cars, experimental studies of the intensity of fleet utilization, time of failure-free operation of vehicles were carried out, and it was demonstrated that the examined transport system is characterized by a satisfactory, stationary readiness coefficient. The developmental possibilities of the presented modelling method were emphasized.

Keywords: *vehicle exploitation system, special vehicles, readiness, semi-Markov model.*

System eksploatacji samochodów, które realizują zadania interwencyjne, niepodlegający systematycznym zmianom może być modelowany jako stacjonarny proces stochastyczny. Taki model pozwala wyznaczyć probabilistyczne wskaźniki bieżącej i granicznej gotowości systemu. W niniejszym artykule, do modelowania systemu eksploatacji pojazdów specjalnych, wykorzystano proces semi-Markowa, oparty na trzech stanach eksploatacyjnych: użytkowania, postoju użytkowego i naprawy. Na przykładzie próby radiowozów policyjnych przeprowadzono doświadczalne badania intensywności użytkowania floty, czasu bezawaryjnej pracy pojazdów a także wykazano, że badany system transportowy charakteryzuje się zadowalającym, stacjonarnym współczynnikiem gotowości. Podkreślono rozwojowe możliwości przedstawionej metody modelowania.

Słowa kluczowe*: system eksploatacji samochodów, pojazdy specjalne, gotowość, model semi-Markowa.*

Introduction

The process of commercial vehicle exploitation can be analysed both in road transport companies, which operate in market conditions, as well as in rescue services and other services responsible for national security, such as the fire brigade, army, police, ambulance service. In the first group, the most important criterion for assessing the operational quality of a vehicle is efficiency, usually measured as a profit/cost ratio [1]. The second group, especially the Police, is identified above all with keeping the peace, protecting the lives and health of people and property, and preserving order. Therefore, most of the research in this thematic area is mainly related to security issues in the broad sense, and concerns, for example:

1. Estimates of the likelihood of a fatal accident when driving a police car [3] and the assessment of the risk of traffic incidents, including serious injuries resulting from participation in police operations [6, 25].
2. Possibilities of increasing the security level of police operations through the use of special methods or devices, such as the bulletproof panels mounted on police cars proposed by Michaelson [27] or the warning light systems described by Lyons [26].
3. Methods for planning and optimising patrol routes [8, 10], with particular emphasis on security issues [4] as well as the necessary number of patrol cars depending on the intensity of the activities carried out and the time of their occurrence [22].

On the other hand, the readiness and reliability of police vehicles is considered a kind of status quo. The studies presented in the literature on the assessment of readiness of complex intervention systems (not only police ones) are of a unitary nature. This is mainly due to the limitations associated with the confidential nature of the empirical data. Record and billing documentation is usually kept in paper form and the practice of creating electronic databases encounters organisational barriers.

Transport tasks are complex processes, which means that their modelling based on classical techniques of reliability theory can be complicated and may not produce satisfactory results [21]. Alternative methods are used in such a case, e.g. reliable phase diagrams proposed by Lu et al. [24] or Dong et al. [9], as well as Markov processes [11, 16, 34], which are particularly popular in the readiness assessment. In the literature one can find models describing single means of transport, e.g. a passenger car – as in the case of Girtler and Ślęzak [12], a bus in the case of Landowski et al. [23], or a helicopter in the case of Szawłowski [23]. Complex transport systems are also studied. Theoretical basis for such considerations are included in the papers [2, 13, 20]. The systems are analysed as a whole [7, 32, 35] or their individual components are considered independently, and each of them is described by a separate model. Often, the authors point to Markov processes as a tool to solve a particular exploitation problem [29, 30]. Unfortunately, transport systems models based on empirical data are few and far between. There are only single studies available,

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

e.g. Migawa [28] studied the city bus exploitation system in this way, Żurek and Tomaszewska [39] analysed aircrafts and Restel [31] analysed urban rail transport systems.

The literature review shows that the Markov models are a good tool to assess the readiness of both whole systems and individual objects [5, 19]. However, they have their own requirements and limitations. These include, first of all, the form of available observations, the distribution of which should be exponential. It is an element that is often omitted in the presented analyses, which causes the use of Markov processes to be abused. It is more difficult to estimate parameters in the case of semi-Markov models, which is why they are less popular. They have less restrictive requirements concerning the form of distributions of studied variables (they can be arbitrary), therefore they are proposed in this article as a tool for police car fleet assessment. The aim of the presented study was to estimate the level of their readiness at the assumption of three operating states: operation, ready-to-be-used and repair (technical maintenance) and to present a method for a stochastic description of the exploitation process. Moreover, the intention of the authors was to emphasize that the three-state exploitation model may be a useful and sufficient tool to assess the readiness of special vehicles. The application of such a model does not require complicated calculations as it is the case with complex multi-state models and can be used in the current practice of fleet management.

2. Exploitation studies and preliminary analysis of the results

The subject of the study were police cars, performing patrol and intervention tasks in the capital city of Warsaw. A total of 20 Kia brand marked passenger cars were analysed. All vehicles came from a single production batch, which allowed the sample to be considered as homogeneous. The source database was the documentation of the use of police cars concerning police patrols and the records of technical services and repairs.

On the basis of the collected observations, a three-element set of operating states $S = \{S_1, S_2, S_3\}$ of the vehicles was singled out:

- operation (S_1),
- ready-to-be-used (S_2),
- repair (including technical maintenance) (S_3).

It was assumed that the time the vehicle remains in the state S_1 (state S_1 duration) falls within the range from the moment of departure in order to perform an intervention task (patrol) to the moment of returning to the depot. The time the vehicle remains in the state S_2 (state S_2 duration) falls within the range from the moment of starting a stop in the depot waiting for the instruction to perform the task until the moment of departure. The time the vehicle remains in the state S_3 (state S_3 duration) is determined by the time when the technical maintenance starts and ends.

Then, based on the actual interstate relations, permitted transitions were determined, which are presented in Fig. 1 in the form of a graph.

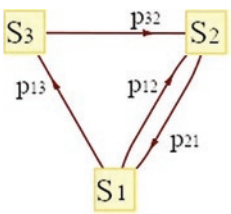


Fig. 1. Permitted transitions graph

The analysis of statistical time distributions (expressed in minutes) of individual operating states was also carried out. Matching of real observations to selected theoretical distributions (normal, log-normal, exponential, Gamma and Weibull) was examined. The parameters of these distributions were estimated using the Statistica program, applying the highest reliability method. The quality of matching was assessed by comparing the distribution of observed frequencies with the expected ones. The statistics of the Kolmogorov-Smirnov test and the Akaike Information Criterion were calculated. On the basis of the results obtained, a gamma distribution was selected as the most suitable one. An exemplary analysis was presented for the distribution of the operation state - S_1 (Fig. 2).

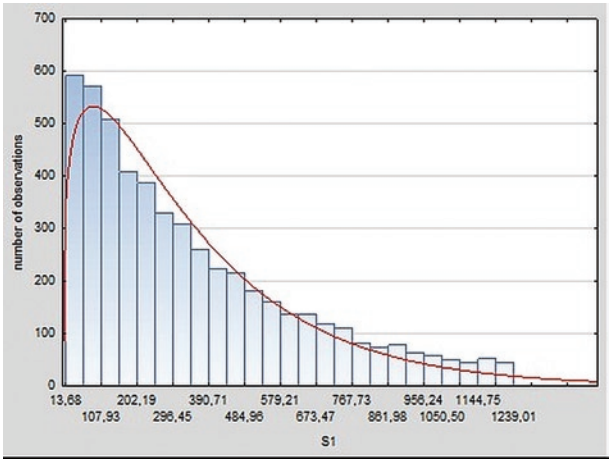


Fig. 2. Histogram of state duration times S_1

3. Estimation of parameters of semi-Markov model

3.1. Basic characteristics

The conclusion from the preliminary analyses was the lack of possibility to use the Markov model (it requires exponential form of distributions of variables) and the assumption to carry out analyses with the use of the semi-Markov model, for which the form of distributions may be arbitrary.

For the examined process of car exploitation, a semi-Markov model with a finite set of states was determined by means of the Markov renewal process, based on [12, 13, 20]:

For N denoting a set of non-negative integers, S – a certain finite set, R_+ – a set of non-negative real numbers, while $\Omega, \mathcal{F}, \mathrm{P})$ – a probabilistic space in which for each $n \in \mathrm{N}$ random variables are specified:

$$\xi_n : \Omega \to S \tag{1}$$

$$\vartheta_n : \Theta \to R_+ \tag{2}$$

Two-dimensional sequence of random variables $\{\xi_n, \vartheta_n : n \in N\}$ is referred to as the Markov renewal process if for each $n \in N$, $i, j \in S$, $t \in R_+$:

$$P\{\xi_{n+1} = j, \vartheta_{n+1} < t | \xi_n = i, \xi_{n-1}, \ldots, \xi_0, \vartheta_n, \ldots \vartheta_0\} = P\{\xi_{n+1} = j, \vartheta_{n+1} < t | \xi_n = i\} \tag{3}$$

and

$$P\{\xi_0 = i, \vartheta_0 = 0\} = P\{\xi_0 = i\} \tag{4}$$

This definition shows that Markov renewal process is a specific case of the two-dimensional Markov process [14]. Transition probabilities of this process depend solely on the discrete value of the coordinate. Markov renewal process $\{\xi_n, \vartheta_n : n \in N\}$ is called homogeneous if probabilities:

$$P\{\xi_{n+1} = j, \vartheta_{n+1} < t | \xi_n = i\} = Q_{ij}(t) \tag{5}$$

do not depend on n.

Functional matrix:

$$Q(t) = [Q_{ij}(t)], i, j \in S \tag{6}$$

is called the renewal kernel. Semi-Markov process is defined basing on the homogeneous Markov renewal process 14.

Let:

$$M(t) = \sup\{m \geq 0 : \tau_m \leq t\} \tag{7}$$

where:

$$\tau_m = \vartheta_0 + \vartheta_1 + \ldots + \vartheta_m \tag{8}$$

The stochastic process $\{M(t) : t \in R_+\}$ is constant within the range $[\tau_m, \tau_{m+1})$. The stochastic process $\{X(t) : t \in R_+\}$ is determined by the formula:

$$X(t) = \xi_{M(t)} \tag{9}$$

is a semi-Markov model,

Defining a model semi-Markov process requires defining, in addition to the kernel of the process, its initial distribution [13, 17, 38]. The process of vehicle exploitation was divided into three phases of random duration. In this case, the renewal kernel of the semi-Markov process, according to the graph of permitted transitions (Fig. 1), takes the form:

$$Q(t) = \begin{bmatrix} 0 & Q_{12}(t) & Q_{13}(t) \\ Q_{21}(t) & 0 & 0 \\ 0 & Q_{31}(t) & 0 \end{bmatrix} \tag{10}$$

This matrix constitutes a model of changes in the distinguished states of the process. The non-zero elements $Q_{ij}(\mathrm{t})$ of the matrix $Q(t)$ are the conditional probabilities of the process of transition from the state S_i to the state S_j, within a time period of not more than t, specified according to the formula (11). They depend on the distribution of random variables, namely the process durations in the distinguished states:

$$Q_{ij}(t) = P\left(X(\tau_{m+1}) = j,\ \tau_{m+1} - \tau_m \leq t | X(\tau_m) = i\right) \text{ for } t \geq 0 \tag{11}$$

where a random variable τ_m means the moment of m -th change of state.

Initial distribution: $p_i(0),\ i \in S = \{1, 2, 3\}$ was adopted in the following form:

$$p_i(0) = \begin{cases} 1, \text{ if } i = 1 \\ 0, \text{ if } i \neq 1 \end{cases} \tag{12}$$

where:

$$p_i(0) = P\{X(0) = i\}, \quad i = 1, 2, 3 \tag{13}$$

These elements make it possible to determine the probabilistic parameters of the exploitation process that are being searched for. For the semi-Markov model, the transition probabilities, defined as conditional probabilities [15], are important:

$$P_{ij}(t) = P\{X(t) = j | X(0) = i\}, i, j \in S \tag{14}$$

$P_{ij}(t)$ are the probabilities of transition from the state S_i to the state S_j at the moment t. They were calculated on the basis of real interstate relations, according to the formula (15).

$$p_{ij} = \frac{n_{ij}}{\sum_{k \in S} n_{ik}} \tag{15}$$

where:

n_{ij} – number of transitions from the state S_i to the state S_j,

$\sum_{k \in S} n_{ik}$ - number of all transitions (exits) from the state S_i,

The distribution of probability of changes of the distinguished operating states (in one step), assuming that each graph arch of the exploitation process representation (Fig. 1), connecting two states of the process, corresponds to the value of probability p_{ij}, is presented in Table 1.

Table 1. Transitions probabilities matrix p_{ij}

p_{ij}	S_1	S_2	S_3
S_1	0	0.8	0.2
S_2	1	0	0
S_3	0	1	0

The calculated values of probabilities of transitions refer to sets of states, not time period. For example, $p_{13} = 0.2$ means that among all the exits from the state S_1, transitions from the state S_1 to S_2 constitute 20%.

3.2. Boundary properties

An important role in the study of the process of exploitation of cars modelled by the Markov chain is played by its boundary properties [13, 20], and especially by the boundaries of probabilities $p_j(n)$ and $p_{ij}(n)$ at $n \to \infty$, , which describe the behaviour of the process after a long time [13, 36]. An important concept in this respect is the stationary distribution of homogeneous Markov chain, described by the vector Π [14]:

$$\Pi = [\pi_1, \pi_2, \pi_3] \tag{16}$$

so as:

$$\Pi = \Pi \mathrm{P} \tag{17}$$

where:

$$\mathrm{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \tag{18}$$

and:

$$\sum_{j=1}^{3} \pi_j = 1 \tag{19}$$

this means that if the chain at a certain point in time m reaches the stationary distribution, then for each subsequent moment n greater than m the unconditional distribution will remain the same.

In the case of the examined process, some limits exist:

$$\lim_{n \to \infty} p_{ij}(n) = \pi_j \qquad i, j = 1,2,3 \tag{20}$$

where:

$p_{ij}(n)$ – probability of transition from the state S_i to the state S_j in n steps.

Calculated probability matrix of changes in operating states inserted into the Markov chain process (Table 1) made it possible to determine the stationary probabilities π_j, according to a system of equations (17).

For the examined process, for the 3-state model, the estimation of the stationary probabilities π_j required the solution of the matrix equation:

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix}^T \cdot \begin{bmatrix} 0 & p_{12} & p_{13} \\ p_{21} & 0 & 0 \\ 0 & p_{32} & 0 \end{bmatrix} = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix}^T \tag{21}$$

with the normalization condition:

$$\pi_1 + \pi_2 + \pi_3 = 1 \tag{22}$$

which is equivalent to the following system of equations:

$$\begin{cases} \pi_2 \cdot p_{21} = \pi_1 \\ \pi_1 \cdot p_{12} + \pi_3 \cdot p_{32} = \pi_2 \\ \pi_1 \cdot p_{13} = \pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \tag{23}$$

After substituting the value of probability of transitions (Table 1), we get:

$$\begin{cases} \pi_2 = \pi_1 \\ 0.8\pi_1 + \pi_3 = \pi_2 \\ 0.2\pi_1 = \pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \tag{24}$$

The solution of the system of equations is presented in Table 2.

Table 2. Stationary probabilities π_j of the distinguished operating states

	S_1	S_2	S_3
π_i	0.455	0.455	0.09
$\pi_i[\%]$	45.5	45.5	9

Next, on the basis of the directed graph (Fig. 1) determining the probability of transitions of Markov chain states (Table 1), and on the basis of empirical times t_{ij} of duration of individual states, conditional estimations of expected $E(T_{ij})$ times of duration of process $X(t)$ states were made on the basis of the estimator defined by the formula (24):

$$\widehat{E(T_{ij})} = \overline{T}_{ij} = \frac{t_{ij}}{\sum_{j \in S} t_{ij}} \tag{25}$$

The matrix $\overline{T} = \left[\overline{T}_{ij}\right]$, $i, j = 1,2,3$ of estimated conditional values of expected times T_{ij} is presented in Table 3.

Table 3. Estimated expected values of conditional times T_{ij}

$\overline{T}_{ij}[minutes]$	S_1	S_2	S_3
S_1		844	845
S_2	479		
S_3		388	

When the elements of the matrix P and $\overline{T}$ are known, the expected values $ET_i, i = 1,2,3$ of the unconditional duration times of individual states of the process can be estimated according to the dependency:

$$\widehat{ET_i} = \overline{T}_i = \sum_{j=1}^{3} p_{ij} \cdot \overline{T}_{ij} \tag{26}$$

For the examined 3-state process of vehicle exploitation, the problem of estimating the values of expected unconditional duration of individual states of the process boiled down to the solution of the following system of equations:

$$\begin{cases} \overline{T}_1 = p_{12} \cdot \overline{T}_{12} + p_{13} \cdot \overline{T}_{13} \\ \overline{T}_2 = p_{21} \cdot \overline{T}_{21} \\ \overline{T}_3 = p_{32} \cdot \overline{T}_{32} \end{cases} \tag{27}$$

The estimated values of unconditional times $\overline{T}_i$ are shown in Table 4.

The random variables $T_i, i = 1,2,3$ have finite positive expected values. This makes it possible to determine the boundary distribution of the semi-Markov process. Based on the stationary distribution of the inserted Markov chain (Table 2) and the estimated expected values

Table 4. Unconditional times $\overline{T}_i$ [minutes] of process duration in 3 operating states

state	$\overline{T}_i$ [minutes]
1	844.2
2	479
3	388

of the process duration times (Table 4), boundary probabilities were estimated according to the formula (28) [20]:

$$P_i = \frac{\pi_i \cdot \overline{T}_i}{\sum_{k \in S} \pi_k \cdot \overline{T}_k}, i = 1, 2, 3 \tag{28}$$

The calculated boundary distribution of probability of semi-Markov process states is presented in Table 5.

Table 5. Boundary probabilities distribution P_i

Percentage	P_1	P_2	P_3
probabilities	0.6026	0.3419	0.0555
distribution	60	34	6

The values P_i constitute boundary probabilities determining that in a long period of operation ($t \to \infty$) the vehicle will remain in a given operating state.

The highest values were achieved for the state of operation (60%), which is a very good result. Ready-to-be-used reaches the boundary value of 34%, which is also a satisfactory result and shows, on the one hand, a high level of readiness of the examined vehicles and, on the other hand, a significant reserve which, however, in the case of structures operating in an unforeseen, intervention-based manner, seems rational. In boundary terms, there is only 5.5% probability of vehicles being a repair state.

The technical readiness factor K is the sum of appropriate probabilities of reliability states. For the proposed model of vehicle operation, the states S_1 and S_2 are roadworthy, while the state S_3 is the state of unfitness. Hence, the readiness of the examined vehicles can be calculated as the sum of the boundary probabilities of the states S_1 and S_2:

$$K = P_1 + P_2 \tag{29}$$

The calculated readiness factor is $K = 94.45$ and means that the vehicles from the examined group for almost 95% of the time remain in the technical readiness state.

3.3. Time of first transition of the vehicle exploitation process to a subset of states (time of failure-free operation)

Another important characteristic describing the processes of vehicle exploitation is the time of the first transition of the process to a separate state or a set of states $\{A\}$ [18]. Based on the distribution of this time and its parameters, the probability of vehicles being in a particular state or set of states may be determined [20, 37]. Function in a form:

$$\Phi_{iA}(t) = P(\Theta_A \le t | X(0) = i), t \ge 0 \tag{30}$$

is a distribution function of the distribution of a random variable $\Theta_A = \tau_{\Delta_A}$, which means the time elapsing from the moment when the semi-Markov process takes the value $i \in A'$ until the moment when the process takes any value from the subset of states A, where $A \subset S$ and $A' = S - A$.
while:

$$\Delta_A = \min\{n \in N : X(\tau_n) \in A\} \tag{31}$$

For regular semi-Markov processes, in which the subset A is strongly achievable from any state belonging to A', random variables T_{ij} have finite and positive expected values $E(T_{ij})$, there are expected values $E(\Theta_{A'})$ and they are the only solutions of the system of equations [13, 20]:

$$(I - P_{A'})\overline{\Theta}_{A'} = T_{A'} \tag{32}$$

where:

$P_{A'}$ – probability matrix of transitions within the set A'

$\overline{\mathrm{E}}_{A'}$ – process kernel specified in the set A'

$T_{A'}$ – random variables of unconditional duration times of the process in the set of states A'

Since in the process under consideration the transport task will be performed if there is no failure of the means of transport, the distribution of time of task execution (failure-free operation of the system) can be found by reducing the original model by the state S_3 - repair. In such a case the subset of states $A' = \{S_1, S_2\}$, while the subset of states $A = \{S_3\}$, and the elements of the equation (32) take the form:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \overline{\Theta}_{A'} = \begin{bmatrix} \overline{\Theta}_{13} \\ \overline{\Theta}_{23} \end{bmatrix}, \quad \overline{T}_{A'} = \begin{bmatrix} E(T_1) \\ E(T_2) \end{bmatrix}, \tag{33}$$

$$P_{A'}(s) = \begin{bmatrix} 0 & p_{12} \\ p_{21} & 0 \end{bmatrix} \tag{34}$$

where:

$$P_{A'} = [p_{ik}] \quad i,k \in A' \tag{35}$$

is a sub matrix of matrix P_{ij} (Table 1). Random variable Θ_{ij} means the time elapsed between the initial time and the time when the repair condition is first reached, provided that one of the conditions of set A' has commenced at the time considered initial. Hence, it means the time of system failure-free operation. For the analysed semi-Markov model, the matrix equation (32) takes the following form:

$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & p_{12} \\ p_{21} & 0 \end{bmatrix}\right) \cdot \begin{bmatrix} \overline{\Theta}_{13} \\ \overline{\Theta}_{23} \end{bmatrix} = \begin{bmatrix} E(T_1) \\ E(T_2) \end{bmatrix} \tag{36}$$

After substituting appropriate values from Table 1 and Table 4 we get:

$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0.8 \\ 1 & 0 \end{bmatrix}\right) \cdot \begin{bmatrix} \overline{\Theta}_{13} \\ \overline{\Theta}_{23} \end{bmatrix} = \begin{bmatrix} 844.2 \\ 479 \end{bmatrix} \tag{37}$$

which comes down to solving the system of equations:

$$\begin{cases} \bar{\Theta}_{13} - 0.8\bar{\Theta}_{23} = 844.2 \\ -\bar{\Theta}_{13} + \bar{\Theta}_{23} = 479 \end{cases} \tag{38}$$

The results of calculations of the above equations are presented in Table 6.

Table 6. Values of the elements of the matrix $\bar{\Theta}$ of the time of first transition for all vehicles

$\bar{\Theta}$	$[min]$	$[h]$
$\bar{\Theta}_{13}$	6137	102.3
$\bar{\Theta}_{13}$	6616	110.3

If the initial decomposition of the exploitation process is a vector:

$$p = [p_1, p_2, p_3] \tag{39}$$

which in the examined process, according to the original assumption (12), takes the form:

$$p = [1,0,0] \tag{40}$$

then the first row of the single-column matrix solving this equation is the expected value of the task execution time, which in this case is more than 102 hours.

It is also possible to determine the time distribution for the correct operation of the object. Using the information that the probability of the transition $P_{ij}(t)$, defined as conditional probabilities [20]:

$$P_{ij}(t) = P\{X(t) = j \mid X(0) = i\},\ i, j \in S \tag{41}$$

fulfil the Feller's equations:

$$P_{ij}(t) = \delta_{ij}[1 - G_i(t)] + \sum_{k \in S} \int_0^t P_{kj}(t-x) dQ_{ik}(x),\ i, j \in S \tag{42}$$

it is possible to find the solution to this system using Laplace – Stieltjes transformation:

$$P_{ij}(t) = \delta_{ij}[1 - G_i(t)] + \sum_{k \in S} \int_0^t P_{kj}(t-x) dQ_{ik}(x),\ i, j \in S \tag{43}$$

$$\tilde{p}_{ij}(s) = \int_0^\infty e^{-st} dP_{ij}(t) \tag{44}$$

$$\tilde{q}_{ik}(s) = \int_0^\infty e^{-st} dQ_{ik}(t) \tag{45}$$

where $Q_{ik}(t)$ is the kernel of the process of renewal of a subset of states A' while $G_i(t)$ denotes the distribution function of a random variable T_i of the duration of the *i-th* state of the semi-Markov process, regardless of the state to which the transition occurs at the moment τ_{n+1} [13]:

$$G_i(t) = P\{T_i < t\} = P\{\tau_{n+1} - \tau_n < t / X(\tau_n) = 1\}, i \in S \tag{46}$$

In this case, the above system of integral equations is correspondent to the system of algebraic equations with unknown transforms $p_{ij}(s),\ i, j \in S$:

$$\tilde{p}_{ij}(s) = \delta_{ij}\left[\frac{1 - \tilde{g}_i(s)}{s}\right] + \sum_{k \in S} \tilde{q}_{ik}(s) \tilde{p}_{kj}(s),\ i, j \in S \tag{47}$$

the system in the matrix notation takes the form:

$$\tilde{P}(s) = \frac{1}{s}[I - \tilde{q}(s)]^{-1}[1 - \tilde{g}(s)] \tag{48}$$

When solved, a transforms matrix is obtained. Since the initial state is the state S_1, the first line is simultaneously a one-dimensional distribution of the process.

For of the examined system:

$$Q(t) = \begin{bmatrix} 0 & Q_{12}(t) \\ Q_{21}(t) & 0 \end{bmatrix} \tag{48}$$

where $Q_{12}(t)$ and $Q_{21}(t)$ are distribution functions of estimated Gamma decompositions:

$$Q_{12}(t) = \frac{e^{-\frac{t}{\beta_1}} t^{-1+\alpha_1} \beta^{-\alpha_1}}{\Gamma[\alpha_1]}, \quad t > 0 \tag{50}$$

$$Q_{21}(t) = \frac{e^{-\frac{t}{\beta_2}} t^{-1+\alpha_2} \beta^{-\alpha_2}}{\Gamma[\alpha_2]} \quad t > 0 \tag{51}$$

and:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} \tag{52}$$

since for the Gamma decomposition the Laplace-Stieltejes transform takes the form:

$$\tilde{f}(s) = \left(\frac{\beta}{\beta + s}\right)^\alpha \tag{53}$$

the elements of equation (48) take the form:

$$\frac{1}{s}[I - \tilde{q}(s)]^{-1} = \begin{bmatrix} \dfrac{0.1}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} & \dfrac{0.1\beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} \\ \dfrac{0.1\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} & \dfrac{0.1}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} \end{bmatrix} \tag{54}$$

and:

$$[1 - \tilde{g}(s)] = \begin{bmatrix} 1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1} & 1 \\ 1 & 1 - \beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1} \end{bmatrix} \tag{55}$$

The solution is a matrix whose elements of the first line are as follows:

$$\tilde{P}_1(s) = \frac{0.1\beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} + \frac{0.1\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\right)}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} \tag{56}$$

$$\tilde{P}_2(s)=\frac{0.1}{s\left(1-\beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)}+\frac{0.1\beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\left(1-\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)}{s\left(1-\beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} \tag{57}$$

After calculating the reverse transforms, the boundary distribution of the intensity of use of the object is obtained. For the state S_1 we get the function in the form:

$$P_1(t)=0.0026857\,\mathrm{e}^{-0.66956t}-0.006151\,\mathrm{e}^{-0.23043t}-0.129546\,\mathrm{e}^{-0.0446957t}+0.6$$

The graph of this function is shown in Fig. 3

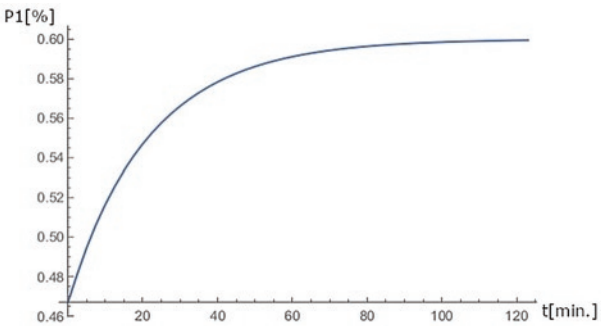


Fig. 3. Function graph $P_1(t)$

The function stabilizes in about 120 minutes, and within the boundary, for $t\to\infty$, it aims towards the previously calculated boundary value of the semi-Markov process of $P_1=60\%$.

4. Conclusions

The use of semi-Markov processes allows to determine the boundary readiness factor and to carry out the analysis of the duration times of distinguished operating states of special vehicles. It also enables an objective assessment of the intensity of vehicle operation and the time of its failure-free operation. Analysing readiness factors, it is possible to search for optimal algorithms of vehicle operation and maintenance, as well as to analyse the quality of vehicle fleet selection.

The validity of the above assumptions was confirmed by the conducted research. The proposed semi-Markov model made it possible to diagnose the system of exploitation of police cars indicating that it is characterized by a satisfactory level of probability of vehicles being in the state of operation ($P_1=0.6$) and ready-to-be-used ($P_2=0.34$). The forecast technical readiness factor amounted to $K=95\%$.

Therefore, the effectiveness of the application of semi-Markov processes to model the readiness of special vehicle exploitation systems has been demonstrated. The three-state model distinguishing the state of operation of the vehicle, the state of ready-to-be-used and the state of repair (technical maintenance) proved to be justified. In this case, it was not necessary to create complex, multi-state structures of the exploitation process model requiring advanced computational programs. The presented, three-state model is expandable in a situation where a deeper analysis of selected aspects of the system readiness would be necessary.

References

1. Andrzejczak K, Młyńczak M, Selech J. Poisson-distributed failures in the predicting of the cost of corrective maintenance. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(4): 602-609, https://doi.org/10.17531/ein.2018.4.11.
2. Bain L.J, Engelhardt M. Introduction to Probability and Mathematical Statistics. Second Edition. California: Cengage Learning, 2000.
3. Becker L.R, Zaloshnja E, Levick N, Guohua L, Miller T. R. Relative risk of injury and death in ambulances and other emergency vehicles. Accident Analysis & Prevention 2003; 35(6): 941-948, https://doi.org/10.1016/S0001-4575(02)00102-1.
4. Behm G.W, Huber W.B, Noll A.J, Pelaez R. A Method and system for safe emergency vehicle operation using route calculation. United States Patent US8842021B2, 2014.
5. Cheng Q, Sun B, Zhao Y, Gu P. A method to analyze the machining accuracy reliability sensitivity of machine tools based on Fast Markov Chain simulation. Podejście do analizy czułości niezawodnościowej dokładności obrabiarek oparte na symulacji metodą szybkich łańcuchów Markowa. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (4): 552-564, https://doi.org/10.17531/ein.2016.4.10.
6. Chu H. C. Risk factors for the severity of injury incurred in crashes involving on-duty police cars. Traffic injury prevention 2016, (5)17: 495-501, https://doi.org/10.1080/15389588.2015.1109082.
7. Dekker R, Nicolai R.P, Kallenberg L.C.M, Maintenance and Markov decision models. In Wiley StatsRef: Statistics Reference Online (eds Balakrishnan N, Colton T, Everitt B, Piegorsch W, Ruggeri F, Teugels J.L.). John Wiley & Sons, 2014, https://doi.org/10.1002/9781118445112.stat03960.
8. Dinc S, Dinc I. Evaluation of Unsupervised Classification on Police Patrol Zone Design Problem. SoutheastCon 2018. St Petersburg, 2018: 1-7, https://doi.org/10.1109/SECON.2018.8478908.
9. Dong W, Liu S, Yang X, Wang H, Fang Z. Balancing reliability and maintenance cost rate of multi-state components with fault interval omission. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 37-45, https://doi.org/10.17531/ein.2019.1.5.
10. Elliott T, Payne A, Atkison T, Smith R. Algorithms in Law Enforcement: Toward Optimal Patrol and Deployment Algorithms. Proceedings of the 2018 International Conference on Information and Knowledge Engineering IKE'18. Las Vegas, 2018: 93-99.
11. Ge H, Tomasevicz C.L, Asgarpoor S. Optimum Maintenance Policy with Inspection by Semi-Markov Decision Processes. 39th North American Power Symposium, Las Cruces, 2007: 541-546, https://doi.org/10.1109/NAPS.2007.4402363.
12. Girtler J, Ślęzak M. Application of the theory of semi-Markov processes to the development of a reliability model of an automotive vehicle. Archiwum Motoryzacji 2012; 2: 15-27, https://doi.org/10.5604/1234754X.1066721.
13. Grabski F. Semi-Markov Processes. Applications in System Reliability and Maintenance. Elsevier, 2015, https://doi.org/10.1016/B978-0-12-800518-7.00004-1.
14. Grabski F. Teoria semi-Markowskich procesów eksploatacji obiektów technicznych. - The theory of semi-Markov processes of technical object exploitation Gdynia: Zeszyty Naukowe Wyższej Szkoły Marynarki Wojennej 75A, 1982.
15. Hong W, Zhou K. A note on the passage time of finite-state Markov chains. Communications in Statistics - Theory and Methods 2017; 46(1): 438-445, https://doi.org/10.1080/03610926.2014.995825.
16. Hu L, Su P, Peng R, Zhang Z. Fuzzy Availability Assessment for Discrete Time Multi-State System under Minor Failures and Repairs by Using Fuzzy Lz-transform. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (2): 179-190, https://doi.org/10.17531/

ein.2017.2.5.
17. Huang X.X, Zou X.L, Guo X.P. A minimization problem of the risk probability in first passage semi-Markov decision processes with loss rates. Science China Mathematics 2015, 58: 1923 - 1938, https://doi.org/10.1007/s11425-015-5029-x.
18. Hunter J.J. The computation of the mean first passage times for Markov chains. Linear Algebra and its Applications 2018; 549: 100-122, https://doi.org/10.1016/j.laa.2018.03.010.
19. Iscioglu F, Kocak A. Dynamic reliability analysis of a multi-state manufacturing system. Eksploatacja i Niezawodnoscć - Maintenance and Reliability 2019; 21 (3): 451-459, https://doi.org/10.17531/ein.2019.3.11.
20. Jaźwiński J, Grabski F. Niektóre problemy modelowania systemów transportowych - Selected problems of transport system modelling. Radom: Instytut Technologii Eksploatacji, 2003.
21. Kaczor G. Modelowanie i ocena niezawodności systemu transportu intermodalnego - Modelling and assessment of the reliability of the intermodal transport system. Logistyka 2015; 3: 2047-2054.
22. Kolesar P.J, Rider K.L, Crabill T.B, Walker W.E. A Queuing-Linear Programming Approach to Scheduling Police Patrol Cars. Operations Research 1975; 23(6):1045-1062, https://doi.org/10.1287/opre.23.6.1045.
23. Landowski B, Muślewski Ł, Knopik L, Bojar P. Semi-Markov model of quality state changes of a selected transport system. Journal of KONES 2017; 24(4): 141-148.
24. Lu J-M, Lundteigen M.A, Liu Y, Wu X-Y. Flexible truncation method for the reliability assessment of phased mission systems with repairable components. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (2): 229-236, https://doi.org/10.17531/ein.2016.2.10.
25. Lundälv J, Philipson Ch, Sarre R. How do we reduce the risk of deaths and injuries from incidents involving police cars? Understanding injury prevention in the Swedish context. Police Practice and Research 2010; 11(5): 437-450, https://doi.org/10.1080/15614263.2010.497333.
26. Lyons H.W. Integrated warning light and rear-view mirror. United States Patent 5851064, 1998.
27. Michaelson E.B. Bulletproof blanket for use with law enforcement vehicles such as police cars. United States Patent 6161462, 2000.
28. Migawa K. Availability control for means of transport in decisive semi-Markov models of exploitation process. Archives of Transport 2012; 4(24): 497-508, https://doi.org/10.2478/v10174-012-0030-4.
29. Młyńczak M. Metodyka badań eksploatacyjnych obiektów mechanicznych - Methodology of exploitation tests of mechanical objects. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, 2012.
30. Muślewski Ł. Control Method for Transport System Operational Quality. Journal of KONES 2009; 3(16): 275-282.
31. Restel F. The Markov reliability and safety model of the railway transportation system. Safety and Reliability: Methodology and Applications - Proceeding of the European Safety and Reliability Conference. London, 2014: 303-311, https://doi.org/10.1201/b17399-46.
32. Świderski A. Inżynieria jakości w wybranych obszarach transportu - Quality engineering in selected areas of transport. Warszawa: Instytut Transportu Samochodowego (Motor Transport Institute), Warszawa 2018.
33. Szawłowski S. Analiza wpływu systemu obsług na gotowość techniczną śmigłowca pokładowego SH-2G - Analysis of the impact of the maintenance system on the technical readiness of the SH-2G ship-based helicopter. Prace Instytutu Lotnictwa 2008; 3-4 (194-195): 326-331.
34. Thomas O.S, Sobanjo J.O. Semi-Markov Decision Process: A Decision Tool for Transportation Infrastructure Management Systems. International Conference on Transportation and Development: Projects and Practices for Prosperity 2016: 384 - 396, https://doi.org/10.1061/9780784479926.036.
35. Woropay M, Żurek J, Migawa K. Model of assessment and shaping of operational readiness of the maintenance subsystem in the transport system. Radom: Instytut Technologii Eksploatacji, 2003.
36. Wu X, Zhang J. Finite approximation of the first passage models for discrete-time Markov decision processes with varying discount factors. Discrete Event Dynamic Systems 2016; 26(4): 669 - 683, https://doi.org/10.1007/s10626-014-0209-3.
37. Wu X, Zou X, Guo X. First passage Markov decision processes with constraints and varying discount factors. Frontiers of Mathematics in China 2015; 10(4): 1005-1023, https://doi.org/10.1007/s11464-015-0479-6.
38. Xie W, Hong Y, Trivedi K. Analysis of a two-level software rejuvenation policy. Reliability Engineering & System Safety 2005; 87: 13-22, https://doi.org/10.1016/j.ress.2004.02.011.
39. Żurek J, Tomaszewska J. Analysis of the exploitation system from the standpoint of readiness, Prace Naukowe Politechniki Warszawskiej, Warszawa 2016; 114: 471 -477.

Anna BORUCKA
Military University of Technology
ul. Kaliskiego 2, 01-476, Warsaw, Poland

Andrzej NIEWCZAS
Motor Transport Institute,
ul. Jagiellońska 80, 03-301, Warsaw, Poland

Kamila HASILOVA
University of Defence,
Kounicova 65, 662 10 Brno, Czech Republic

E-mails: anna.borucka@wat.edu.pl, andrzej.niewczas@its.waw.pl, kamila.hasilova@unob.cz

Article citation info:
ŻYWICA G, KACZMARCZYK TZ. Experimental evaluation of the dynamic properties of an energy microturbine with defects in the rotating system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 670–678, http://dx.doi.org/10.17531/ein.2019.4.17.

Grzegorz ŻYWICA
Tomasz Z. KACZMARCZYK

EXPERIMENTAL EVALUATION OF THE DYNAMIC PROPERTIES OF AN ENERGY MICROTURBINE WITH DEFECTS IN THE ROTATING SYSTEM

EKSPERYMENTALNA OCENA WŁAŚCIWOŚCI DYNAMICZNYCH MIKROTURBINY ENERGETYCZNEJ W OBECNOŚCI DEFEKTÓW UKŁADU WIRUJĄCEGO*

Today's energy systems increasingly use various types of microturbines to produce electricity. A specific feature of such machines is a high-speed rotor, whose rotational speed can be higher than 100,000 rpm. Failure-free operation of high-speed microturbine rotors requires both special design and high precision during the manufacturing process. What is more, proper procedures must be followed during run-up and coast-down phases; and also, dedicated diagnostic systems have to be used. This article discusses the experimental research conducted on a 2.5 kW vapour microturbine that operated in a prototypical combined heat and power plant. A series of measurements was carried out to evaluate the dynamic performance of the machine during normal operation. After the appearance of certain defects in the rotating system, it was necessary to perform a new series of measurements in order to assess the dynamic properties of the machine. The measurements results obtained in the form of vibration velocity spectrums made it possible to define diagnostic symptoms corresponding to particular defects. Similar diagnostic symptoms can occur during the operation of this class of turbomachines.

Keywords: *microturbines, high-speed rotors, damage in a rotating system, vibrations of turbomachines, rotor dynamics.*

We współczesnych systemach energetycznych coraz częściej do wytwarzania energii elektrycznej stosowane są różnego typu mikroturbiny. Charakterystyczną cechą takich maszyn są wysokoobrotowe wirniki, których prędkości obrotowe mogą przekraczać nawet 100 000 obr/min. Praca wirnika w takich warunkach wymaga zastosowania specjalnych rozwiązań konstrukcyjnych i bardzo dużej precyzji wykonania, a podczas eksploatacji zachowania odpowiednich procedur przy rozruchu i odstawieniu, a także stosowania dedykowanych systemów diagnostycznych. W niniejszym artykule zostały omówione badania eksperymentalne mikroturbiny parowej o mocy 2,5 kW, pracującej w prototypowym układzie kogeneracyjnym. Wykonane pomiary obejmowały ocenę stanu dynamicznego podczas normalnej pracy maszyny oraz badania jej właściwości dynamicznych w obecności defektów układu wirującego. Uzyskane wyniki pomiarów, w postaci rozkładów częstotliwościowych drgań, pozwalają na zdefiniowanie symptomów diagnostycznych typowych dla różnych defektów, które mogą pojawić się podczas eksploatacji tej klasy maszyn wirnikowych.

Słowa kluczowe: *mikroturbiny, wysokoobrotowe wirniki, uszkodzenia układu wirującego, drgania maszyn wirnikowych, dynamika wirników.*

1. Introduction

In modern electric power systems, distributed energy sources are playing an increasingly important role because these systems allow to efficiently produce thermal and electrical energy on a small scale from locally-available resources [3]. Depending on the available source of primary or renewable energy and energy demand, different types of thermal microturbines can be used to generate electricity [21]. Their basic features are as follows: power can be adapted to meet customer requirements, high operational availability (they can be switched on and off quickly compared to high-power steam turbines), small dimensions, high-speed rotors, high mobility, relatively low capital and operating costs [11]. All of these features are causing microturbines to be increasingly used, making a major contribution to the development of micro-power cogeneration systems [12].

The term 'microturbine' generally refers to a turbine whose electric power does not exceed 1 MW. Microturbines can be divided into several groups depending on their principle of operation, fluid flow system design, power capacity, fuel type, or application. With regard to thermal microturbines, the main division is into gas and vapour microturbines. The rotors of gas microturbines are powered by exhaust gases from the combustor [2, 21]. Natural gas, biogas or kerosene can be used as fuel in gas microturbines. As for vapour microturbines, their rotors are powered by the vapour of the working medium, which circulates in a closed cycle and is heated using an external heat source [12]. In such a system, the following types of heat sources can be applied: gas or biomass boiler, geothermal source, solar collectors and waste heat from various production processes [7, 18]. The highest efficiency and profitability of installations with gas or vapour microturbines can be reached when the cogeneration is running (that is when there is a demand for both electric energy and thermal energy). Vapour microturbines have lower operating temperatures than gas microturbines and can produce electricity from low-temperature heat sources. Unlike gas mi-

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

croturbines, the operation of vapour microturbines does not depend on the fuel used because they are powered by working mediums that are not present in the combustion chamber. Steam turbines use water as their working medium. For vapour microturbines, various low-boiling mediums [6] are used. A thermodynamic cycle with a low-boiling medium is called the organic Rankine cycle (ORC).

High-speed rotors are used in energy microturbines. Shafts that rotate at high rotational speeds (which may exceed 100,000 rpm [13]) need to be supported by advanced and unconventional bearings such as magnetic [15], foil [28] and gas [19] bearings or even bearings lubricated with low-boiling mediums [14]. In addition to a good dynamic performance of the rotor over a wide rotational speed range, the bearing systems must be of great durability and reliability [7], allowing microturbines to operate without constant technical supervision. With the compact dimensions and low power output of microturbines, no strict procedures have to be followed every time they are set in motion as opposed to big power-station turbines [5]. As microturbine blades are small and rigid, they do not suffer from dynamic problems generally encountered in turbomachines used in the power industry [16]. However, microturbines can cause other dynamic problems that can result, among other things, from high rotational speeds, a wide range of operating speeds, lightness and low stiffness of the casings or their supporting structures being light and not rigid enough. Some low-boiling mediums used in ORC systems may exert a negative effect on microturbine subassemblies if the construction materials used (for example, plastics or elastomers) are not chemically resistant to these mediums. Since low-boiling working mediums are not so widely used as water, microturbine designers do not always have sufficient knowledge of their chemical compatibility with various materials.

Energy microturbines have only become popular in the past few years. In the past, they were scarce and only a few large international companies were involved in their production and research. Therefore, the scientific literature contains little information on construction details or guidelines for the operation of this type of machines. In fact, the results of research on the dynamic properties of microturbines operating under various conditions are almost nowhere to be found. Whereas, for those people who are engaged in the operation and maintenance of fluid-flow machines, such results are a valuable source of information that can be useful in developing a reliable diagnostic system, establishing alert and warning thresholds or for scheduling periodical inspections. The results of vibration measurements of fluid-flow machines make it possible to make structural changes to these machines or their support systems, which can lead to an improvement in their dynamic performance.

It is quite common for large steam turbines to be the subject of research for both scientists and engineers who are involved in the construction and operation of turbomachines. Therefore, there are many articles that present the results of research carried out not only on properly functioning machines but also on malfunctioning ones, where various types of defects are analysed. The experimental research of a 15 MW auxiliary steam turbine is presented in paper [22]. In this case, fractures occurred in the lacing wires that were mounted on the blades to reduce the displacement during their vibration. These fractures were observed in one of the stages located in the low-pressure part of the turbine. Both vibration measurements and experimental modal analysis of the blade were performed to identify the causes of the failure. The numerical analysis was performed as well. Its results are proof that there is a place on the blade in which the stress concentration occurred, which led to the failure. Article [27] also discusses a blade failure that occurred in the low-pressure stage of a 310 MW steam turbine. In this case, stress-corrosion cracking caused the failure. The authors analysed the microstructure of the blade material and concluded that some blades had not been properly tempered. An investigation of the rotor failure of a 60 MW turbine installed in a thermal power plant is discussed in article [1]. A fatigue crack of the shaft appeared after 10 years of operation. The authors of the article applied various diagnostic techniques to detect it, but the main focus was on investigating the fatigue fracture and performing material analyses. Paper [17] discusses torsional vibration and fatigue wear of the shaft of a 660 MW steam turbine. Two methods were considered to repair abrasion defects. A stress concentration occurred in the place where the wear defects are the most visible. An analysis of the impact of temperature changes on fatigue failures of the rotor of a steam turbine with a power of 1000 W is presented in article [24]. Poursaeidi et al. [23] determined temperature distribution on the gas turbine's casing and then analysed its effect on distortions and stress concentrations. An analysis, using an FEM model, was performed for the following power levels: 82, 87 and 96 MW. The results of temperature measurements performed on the real object were used during the creation of the model. Cracks that occurred in the immediate vicinity of certain holes in the casing were related to thermal stress concentrations. Article [25] exemplifies the practical application of the finite element method in an analysis of the causes of cracks in the casing of a helicopter engine. Mechanical and heat loads were taken into account in the analysis. In this case, cracks also occurred in places where the greatest concentrations of thermal stress were observed. In the opinion of the authors, the rotor vibrations had a negative effect on the fatigue life of the casing. The results of the vibration measurements of four large turbomachines (namely, centrifugal fans), both in steady and unsteady states, are presented in paper [4]. The research was carried out during a regular operation, which made it possible to identify the causes of the increased vibration level. The excessive vibrations were due, among other things, to an excessive unbalance of the rotating system, and also due to the fact that the inlet valves were not adjusted correctly. In the case of a large steam turbine, it would be reckless to conduct experimental research on a rotating system if we are fully aware that at least one of its elements is damaged because, under such conditions, a major failure could occur, which could pose a threat to people's lives and health. Therefore, the scientific literature lacks articles that present results of such types of experiments. In fact, various types of numerical models are commonly used to perform failure analyses of large turbines.

As far as the articles on energy microturbines are concerned, most of the experimental research presented in them mainly concern measurements of power, rotational speed and various thermodynamic parameters. Much less attention has been paid to vibration measurements and the study of different types of defects that may occur during operation. Conference paper [10] discusses a diagnostic system applied to a cogeneration system with a 95 kW gas microturbine. The diagnostic system was able to detect defects of the cogeneration system and malfunctions of the measuring system based on the values of the following parameters: thermal/electric power, thermodynamic parameters at the microturbine outlet and rotational speed of the rotor (which reached 70,000 rpm). However, no vibrodiagnostic parameters were measured by this system. Article [26] presents studies related to the dynamic characteristics of the high-speed rotor (30,000 rpm) of a microturbine with a complex geometry of the blade system. It was designed to be applied in a 100 kW gas microturbine. A test rig was built to perform experiments on the rotor. Ceramic rolling bearings were mounted on the test rig to support the rotor. Besides the results of the experimental research, the article presents a numerical model of the rotor, into which the substitute stiffness coefficients of the bearings were incorporated. The model allowed to perform a thorough analysis of the dynamic and thermal properties. The rotor vibration at low speeds was very low due to the high critical speed, which translated into the reliable operation of the machine. Hong et al. presented their research on the dynamic performance of the rotor of a 500 W gas microturbine, whose nominal speed is 100,000 rpm [8]. Numerical and experimental studies were conducted under different levels of the rotor unbalance. Additionally, a numerical model was developed,

which was used to perform a forced vibration analysis of the rotor.

The results of experimental studies on energy microturbines, which have been published so far in the literature, very rarely relate to vibration measurements. The authors of this article have not found any publications that present results of vibration measurements of this type of turbomachines, obtained in the event of malfunction or operation after some defects were detected. As a matter of fact, microturbines offer many possibilities in terms of active diagnostic experiments. Their downtimes do not lead to significant financial losses and, thanks to low production costs, they can undergo various modifications and the fast replacement of parts is not a problem. Although their rotors spin at high speeds, their low mass and compact dimensions make it possible to apply totally reliable protections to ensure the safety of experiments. Compared to large energy turbines, microturbines allow for many possibilities to conduct research (also where there are different defects).

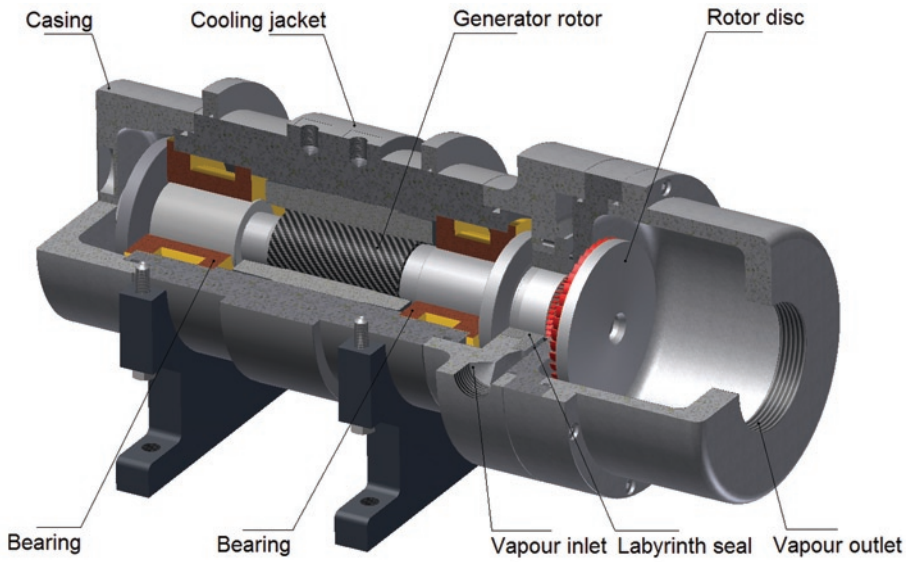


Fig. 1. The sectional view of the microturbine with its marked main parts

The further part of the article focuses on the diagnostic research of a 2.5 kW ORC vapour microturbine. Vibration measurements were carried out in the presence of various types of defects in the rotating system, which occurred during operation. Even though vibrodiagnostic studies of this microturbine have been already presented in article [9], neither the dynamic problems nor the symptoms indicating the appearance of defects were detected at that time. Therefore, the results of previous studies should be treated as basis results, which can be referred to in case of changes in vibration characteristics. Damage, which occurred after the microturbine had been in operation for some time, has brought about new operating characteristics that can be analysed and different types of defects can be detected at the early stages of their development. This opens up new research possibilities in the field of maintenance and damage prevention when the machine is in service.

2. Characteristics of the microturbine

The tested microturbine was developed to produce electric energy in a micro-power ORC cogeneration system [29]. This is a prototypical machine designed within the framework of project No. POIG.01.01.02-00-016/08 (at the Institute of Fluid Flow Machinery of the Polish Academy of Sciences in Gdańsk), in cooperation with the Institute of Turbomachinery, Lodz University of Technology. The maximum thermal and electrical power of the cogeneration system was adapted to meet the average demand of residents of single-family houses. By using a boiler with a thermal power of approximately 25 kW, the microturbine is able to produce electric power of up to 2.5 kW. The microturbine with a generator is small in size due to its target use. The casing has a length of about 350 mm and an outer diameter of less than 200 mm. The sectional view of the microturbine, with its marked main parts, is shown in Fig. 1. In this figure, we can see how the rotor is placed inside the casing and how the bearings are arranged. The shaft is supported by two radial-axial gas bearings, constantly powered by the vapour of HFE-7100, which is the low-boiling working medium that also powers the fluid flow system of the microturbine. Thanks to the use of gas bearings, oil is not necessary for their lubrication and the rotor can reach very high rotational speeds with minimal friction losses in the bearing nodes. Since the bearing lubricant comes from the ORC, no additional lubrication system is required. There is also no possibility for the oil to mix with the low-boiling medium. Microturbines that have this type of design are called oil-free microturbines.

The microturbine's fluid flow system consists of four turbine stages (two centripetal and two centrifugal). Thanks to this, it was possible to minimise the axial force that acts on the thrust bearings. The pressure and temperature of the low-boiling medium at the microturbine inlet is about 11 bar and 180°C, respectively. The nominal rotational speed of the rotor is 24,000 rpm; however, this machine is also adapted to constant operation at lower speeds. The sleeve of the synchronous generator is placed on the shaft of the microturbine, between the journals of the bearings. The generator stator is placed inside the casing, which is equipped with a cooling jacket. A non-contact labyrinth seal separates the microturbine's blade system from the generator and bearings. Since the same working medium is present in both parts of the casing, it was not necessary to use hermetic rotary seals. In this case, minor leaks inside the casing are acceptable and any surplus of the low-boiling liquid is removed by a system of holes. A photo of the entire microturbine, mounted on the test rig, is shown in Fig. 2. Compared to the model presented in Fig. 1, we can see additional rings connected by three threaded rods, which are used as an

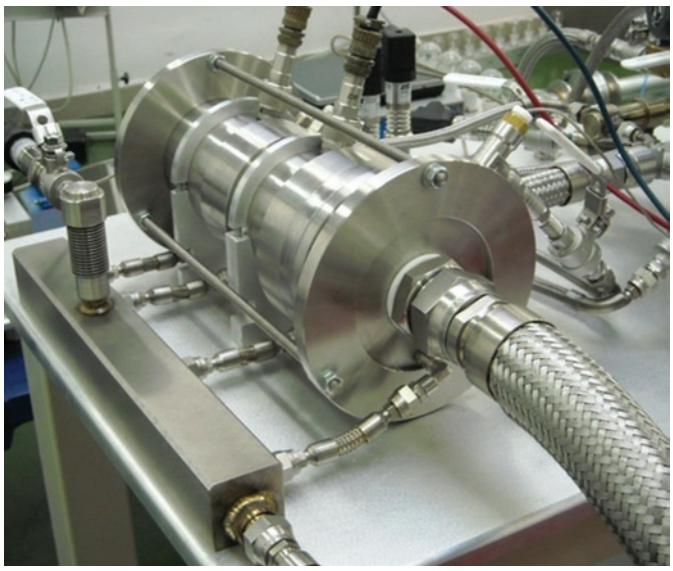


Fig. 2. The microturbine mounted on the test rig

extra measure of protection. The rotor was fitted inside the casing in the axial direction, before mounting the casing covers, the rotor disc and the thrust bearing keep plate. The microturbine was firstly tested using compressed air as a working medium and then tested in the ORC system where a low-boiling medium was present. It was possible, among other things, to check the correct operation of the rotating system, test the measuring system and check the casing for leakage.

Different types of internal and external loads act on particular elements of the microturbine during its operation. The centrifugal force, which results from the high rotational speed, acts on the entire rotating system (including the shaft, the rotor of the generator and the microturbine rotor disc). The residual unbalance of these components causes the rotor to vibrate in the radial and axial direction. Vibrations of the rotor may also be induced by natural vibrations of the rotor disc and blades. When the working medium's hot vapour comes into contact with the microturbine blades, it heats the rotor disc, shaft and casing. Vapour is also delivered to the gas bearings and contributes to the creation of thermal stresses. The vapour flowing through the blade system also causes vibrations of the entire shaft in the radial and axial direction. The operation of the electric generator inherently leads to heat accumulation in the confined space, which results in an increase in the temperature of the shaft, the casing and the generator alone. Since the operating temperature of the generator should not be too high, a water jacket is used for cooling. However, this causes an increase in the temperature gradient and additionally increases thermal stresses in the casing. In addition, these changes in temperature cause thermal deformations of the rotor and casing, which can lead to geometrical incompatibilities and leaks. Electric loads, resulting from the appearance of electromagnetic forces, act on the generator as well. All the loads of the rotating shaft are transmitted to the bearings by the journals. To operate properly, the microturbine's gas bearings require constant lubrication with the low-boiling medium's vapour. In fact, fluid friction in the bearings not only reduces friction losses but also prevents damage to the sliding surfaces. It is of utmost importance that bearings ensure stable operation of the rotor; at the same time, the vibration level should be as low as possible. This is particularly important in the case of high rotational speeds where rotor loads are higher and the lubricant flow can cause unstable bearing operation. All parts located inside the casing are exposed to the low-boiling medium, which is very penetrating and can have a destructive effect on some materials.

According to the above considerations, the microturbine components must be resistant to harsh operating conditions. Therefore, machines of this type require the application of advanced diagnostic methods and systems, which allow constant monitoring of their operating parameters. The results of previous numerical simulations show that changes in some parameters of the microturbine's rotating system discussed here can have a huge impact on the dynamic properties [30]. The same conclusion can be drawn with respect to the experimental research presented in the next part of the article.

3. Experimental studies of vibrations of the microturbine

3.1. Studies of the defect-free microturbine

Experimental studies of the microturbine were carried out in the IMP PAN laboratory, on the test rig that makes it possible to simulate real operating conditions. During the studies, the microturbine blade system and bearings were powered with vapour from the working medium (HFE-7100), at a maximum pressure and temperature of about 11 bar and 180°C, respectively. The values of these parameters are typical operating conditions for micro-power ORC cogeneration systems with multi-fuel boilers. In addition to the vibrodiagnostic measurements, the following operating parameters were measured: thermodynamic parameters of the working medium, rotational speed of the rotor, electric power of the generator. For vibration measurements, a portable vibration analyzer DIAMOND401A (XT version, produced by MBJ Electronics) and a uniaxial accelerometer (model 622B01, produced by PCB Piezotronics) were used. The accelerometer was mounted on the microturbine casing (using a magnetic holder), near the bearing located between the rotor disc and the generator (Fig. 1). Since the casing is made of stainless steel, sensors were fixed using additional stands made of ferromagnetic steel. The vibration level was measured at various locations on the casing in order to choose the appropriate location for the measuring point. The highest vibration level was recorded in the vertical direction, in the vicinity of the bearing situated near the rotor disc, and was considered the most relevant. A computer workstation with the MBJLab software (designed to allow users to diagnose the vibration of various types of machines, including rotating machines) was used for the analysis and visualisation of the measured results.

The results of vibration measurements were presented in the form of vibration velocity spectrums, which greatly facilitated their interpretation. As the maximum rotational speed of the tested machine was as high as 24,000 rpm (400 Hz), the measurements were made in a relatively wide frequency range, from 1 Hz to 800 Hz. A resolution of 1 Hz was used, which is the highest possible resolution of the vibration analyzer mentioned above. The spectral graphs presented in the article are the average results of three consecutive measurements, calculated using the RMS algorithm. To facilitate the comparison of the measured results with each other, the same maximum value (1 mm/s) occurs on the axes of ordinates (representing the RMS vibration velocity amplitude) on all spectral graphs presented here. All the vibration velocity spectrums of the casing of the microturbine were obtained during its stable operation, in other words – at constant power and at a constant rotational speed of the rotor. The performance of measurements at changing speeds would require the application of special signal analysis methods, used for measuring nonstationary states [20]. Moreover, if various measurement results were obtained under varying operating conditions, their comparison among themselves would be difficult.

The results presented in this part of the article were obtained during one of the first tests of the microturbine, which took place several hours after the first start-up. It can be stated that it was a brand new machine without any defects. The vibration of the machine was constantly monitored during these experiments. However, because of the page limit of the article, only selected measurement results (obtained at different rotational speeds) are presented. During these measurements, the rotational speed, power as well as thermodynamic parameters of the working medium's vapour were stable. Fig. 3 shows the vibration velocity spectrum obtained at the rotor speed of 9,300 rpm, while Fig. 4 and Fig. 5 demonstrate vibration velocity spectrums obtained at speeds of 18,060 rpm and 21,060 rpm, respectively.

On all the spectral graphs presented, we can observe one dominant vibration component, occurring at the frequency that corresponds to the rotational frequency of the rotor (the so-called synchronous vibrations). In the consecutive figures, these frequencies are as follows: 155 Hz, 301 Hz and 351 Hz. The vibration velocity amplitude increased as the rotational speed of the rotor increased; it was 0.20 mm/s, 0.36 mm/s and 0.52 mm/s. The constant increase in the vibration level is natural in the speed range analysed as it results from an increase in the centrifugal force (coming from the residual unbalance), which forces rotor vibrations in the transverse direction. The rotor of the tested microturbine is subcritical throughout the operating speed range because the resonant speed (corresponding to the lowest resonant frequency of the bending vibrations), obtained from numerical simulations, was about 130,000 rpm [30]. As a result, a gradual increase in the vibration level as the rotational speed increased was in accordance with our expectations and did not indicate a dynamic problem. In the re-

maining frequency range, there was also an increase in the vibration level at the lowest frequencies (up to about 20 Hz). However, these vibrations were not related to the operation of the microturbine and occurred even at the time when the rotor was not working. In the very low frequency range, vibrations of the construction of the test rig and of the foundation on which it was placed occurred (the floor is located in the laboratory on the first floor, whose ceiling is a lattice construction). The vibrations of these elements were induced, among other things, by the movement of people present in the laboratory and we were unable to eliminate them during the experiments. In addition, at rotational speeds of 18,060 rpm and 21,060 rpm, a slightly higher vibration level occurred at frequencies two times higher and two times lower than the rotational frequency; these are the components 2X and 1/2X. However, due to the very low vibration velocity amplitudes (lower than 0.05 mm/s), these components did not have a significant impact on the dynamic state of the machine.

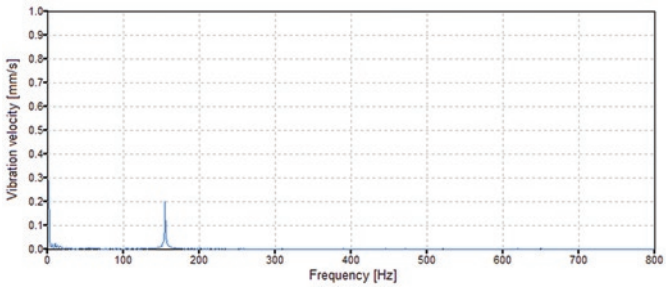


Fig. 3. The vibration velocity spectrum measured on the microturbine casing at a rotational speed of 9,300 rpm (155 Hz)

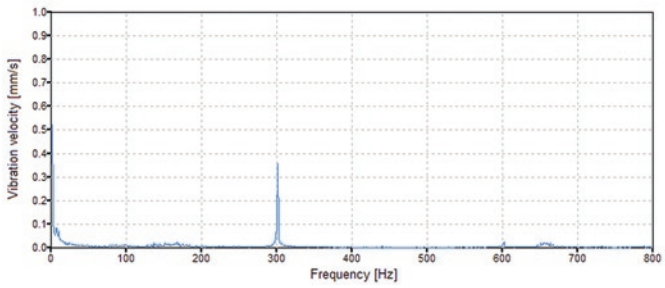


Fig. 4. The vibration velocity spectrum measured on the microturbine casing at a rotational speed of 18,060 rpm (301 Hz)

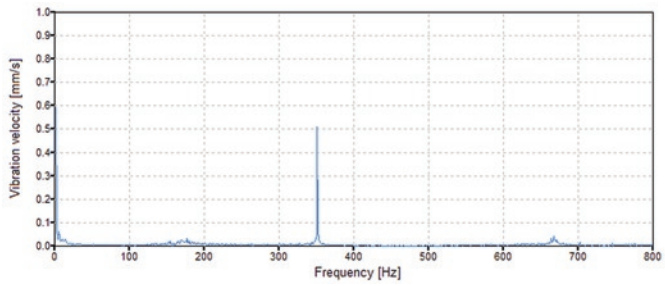


Fig. 5. The vibration velocity spectrum measured on the microturbine casing at a rotational speed of 21,060 rpm (351 Hz)

By evaluating the recorded vibration levels according to the relevant standards, we can affirm that the dynamic state of the tested microturbine was very good. According to the ISO 10816 standard, the overall vibration velocity amplitude (Vrms), measured on the casing of rotating machines with a nominal power up to 15 kW, should not exceed a value of 1.8 mm/s. If this requirement is met, a long-lasting operation is possible without any limitations. As for the machine discussed, the highest vibration velocity amplitude was only 0.52 mm/s and was recorded at a speed of 21,060 rpm. Although the overall vibration level was slightly higher (0.71 mm/s), it was still lower than the maximum value allowed for newly commissioned machines.

Since the tested microturbine worked well across the full range of loads and rotational speeds, the results presented in this part of the article can be treated as reference values. They were obtained from the turbomachine whose technical condition did not raise any doubts. The results of these measurements can, therefore, serve as a reference point for further research or evaluations of the dynamic state of the machine after a certain period of operation. The vibration velocity spectrums obtained can also be useful for determining the symptoms of different types of defects.

3.2. Studies of the microturbine with an unbalanced rotor

The microturbine, after initial tests that proved its proper design parameters and its very good dynamic state, was subjected to further laboratory testing. The purpose of these tests was, among other things, to determine its operating characteristics (such as the output power or the temperature and pressure drops of the working medium). Vibration measurements were also performed. The purpose of the vibrodiagnostic research was to detect potential dynamic problems of the microturbine and the different types of defects that may have occurred due to the experimental nature of the ORC cogeneration system and also due to the fact that it had been tested under different conditions, even the rapidly changing ones [29].

The first symptoms, which indicated the deterioration of the dynamic performance of the machine, were observed after less than 100 hours of operation. It was disturbing to see an increase in the vibration level of the microturbine casing, which was particularly evident at high rotational speeds. A detailed analysis of recorded vibration velocity spectrums showed that apart from the synchronous component (1X), there were no other components on the graphs, however, the level of synchronous vibrations increased significantly. The vibration velocity amplitude was 0.74 mm/s at a speed of 20,520 rpm (Fig. 6) and increased to 0.87 mm/s at a speed of 21,000 rpm (Fig. 7). Compared to the reference value, the vibration level increased by about 67% at a similar rotational speed. Apart from the increased vibration level, there were no other signs that suggested that the microturbine malfunctioned.

To uncover the reason behind such an increase in the vibration level, a decision was made to disassemble the microturbine and conduct a visual inspection of its parts. Already at the beginning of the disassembly, after removing the cover of the casing at the non-drive end (NDE), it was found that the keep plate of the thrust bearing was covered with impurities, as shown in Fig. 8. The cavities in the keep plate that are visible on the photo are not a sign of damage as they were made intentionally when balancing the rotor. When continuing the disassembly of the microturbine, it turned out that a large part of the surface of the rotor was covered with impurities, which caused, among other things, an increase in its unbalance. The increased amplitude of the synchronous component (1X), observed during the vibration measurements, was a typical sign which indicated that the unbalance of the rotating system was too big.

To find the cause of the appearance of impurities on the rotor, the previous tests, which were carried out on the same test rig, were analysed. As this is a general-purpose test rig, it was previously used to study various subassemblies of the ORC system, including different types of circulation pumps. It turned out that the working medium of the ORC system was found to be contaminated with oil that got into the cycle as a result of the failure of one of the pumps tested. This was the membrane pump whose membranes (made of plastic) were not sufficiently resistant to long-lasting contact with the low-boiling medium. Due to the rupture of the membrane, oil from the the pump was flushed out to the ORC. Although the low-boiling medium was replaced in the ORC system after this incident, a significant amount

of oil remained inside the pipeline and other parts of the test rig (for example, in heat exchangers). In later tests, the oil mixed with the low-boiling medium entered the microturbine. When the microturbine ran at high temperatures, oil settled on its components. The surface of the rotor was affected by prolonged contact with the mixture of low-boiling medium and oil. After taking a look at Fig. 8, it becomes clear which part of the casing has been in contact with this mixture.

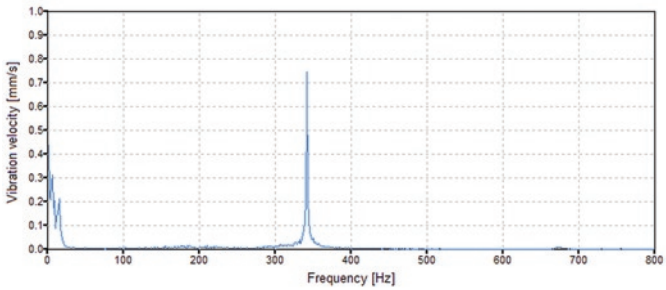


Fig. 6. The vibration velocity spectrum of the microturbine with an unbalanced rotor, measured on the casing at a rotational speed of 20,520 rpm (342 Hz)

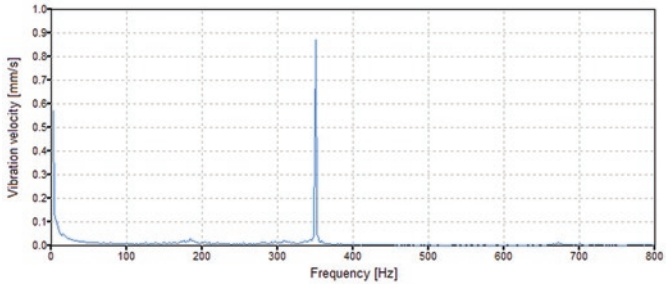


Fig. 7. The vibration velocity spectrum of the microturbine with an unbalanced rotor, measured on the casing at a rotational speed of 21,000 rpm (350 Hz)

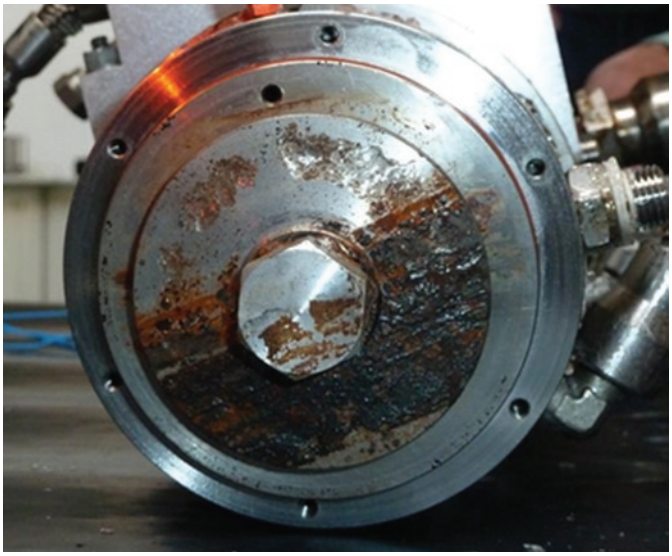


Fig. 8. The bearing keep plate with the defect causing the rotor unbalance

After the disassembly of the microturbine subassemblies, the parts underwent cleaning to remove oil deposits and impurities. The microturbine was restored to its initial technical condition. Prior to further testing, the entire ORC installation was washed and the low-boiling medium was filtered. The actions undertaken have made it possible to re-launch the test rig with the microturbine. Vibration measurements were conducted and their results confirmed the very good dynamic state of the machine. To sum up this part of the article, it can be added that the deterioration of the dynamic state of the turbomachine under test was neither caused by a design mistake nor improper fabrication or assembly but was the result exclusively of inadequate operating conditions, namely, the microturbine was powered using the working medium mixed with oil.

3.3. Research on the microturbine with a malfunctioning bearing

The removal of the remaining oil from the microturbine and the ORC installation restored the test rig to its original state, thus allowing further investigation. Next signs of the microturbine malfunction appeared after one of the several-week-long idle periods. The microturbine was able to operate correctly but only at very low rotational speeds (up to about 6,000 rpm). At higher speeds, it was observed that apart from the 1X component, harmonic components such as 2X, 3X, 4X (and higher) appeared on the vibration velocity spectrums recorded on the casing. In addition, each of these components was not only present at one frequency value but also at the neighbouring ones. The amplitude of the 2X component was similar to that of the 1X component and the amplitudes of subsequent harmonics decreased gradually, as shown in Fig. 9. Despite the disturbing increase in the vibration level, the research was continued, increasing the rotor speed. At rotational speeds above 12,000 rpm, higher harmonics began to gradually disappear but the main component of the spectrum (1X), whose frequency corresponded to the rotational speed of the rotor, also began to appear at the neighbouring frequencies. The vibration spectrum became chaotic, as shown in Fig. 10. These types of spectrums are typical for rotating systems in which there is physical contact between rotating components and the casing or other non-rotating components. Since the vibration characteristic did not improve after several subsequent measurements, it was decided to disassemble the machine and check the technical condition of all its components.

During the disassembly of the microturbine, both the front and rear cover of the casing was removed, and the rotor was taken out. Visual inspection of the components revealed clear signs of wear of one of the gas thrust bearings (Fig. 11), indicating that it malfunctioned. This bearing was located at the free end of the shaft (see Fig. 1). Nearly half of the sliding surface showed signs of friction, which were the result of physical contact of the bearing with the surface of the keep plate. On closer inspection, it was found that 3 out of 10 supply holes were clogged. These holes are used to deliver the lubricating medium (in the form of low-boiling medium vapour) into the space between the sliding surfaces. Since the holes are small in diameter (only 0.4 mm), small particles of different sediments, from different parts of the installation, clogged them. The long downtime was detrimental to the microturbine because impurities accumulated in the supply chambers of the gas bearings. Due to the clogging of the holes, the amount of vapour delivered to the bearing lubrication gap decreased, resulting in a decrease in capacity. The axial force, transmitted by the shaft to the thrust bearing, increased as the flow rate and pressure of the vapour delivered to the microturbine blade system increased. This force was too high for the defective bearing, which resulted in the physical contact between two mating bearing surfaces. Under normal operating conditions, these surfaces are entirely separated by a lubricating film.

A visual inspection of all microturbine components showed that only the sliding surface of one of the thrust bearings was damaged. Since the keep plate was made of a much harder material, the damage only occurred on the surface of the bearing sleeve (made of bronze). Constant monitoring of the vibrations enabled early detection of this problem, so it was easy to repair the damage; the entire surface was smoothed by grinding. After the repair, the operation of the microturbine seemed normal and both the vibration velocity spectrum and the overall vibration level differed only slightly from the reference values.

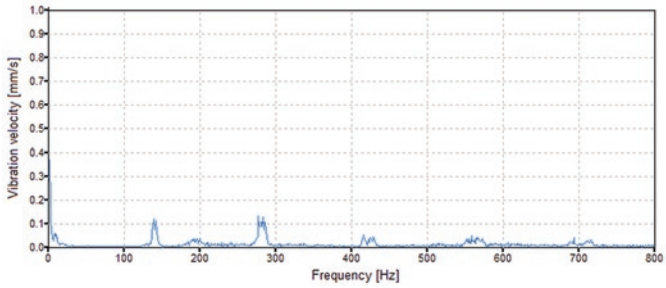


Fig. 9. The vibration velocity spectrum of the microturbine with a damaged thrust bearing, measured on the casing at a rotational speed of 8,340 rpm (139 Hz)

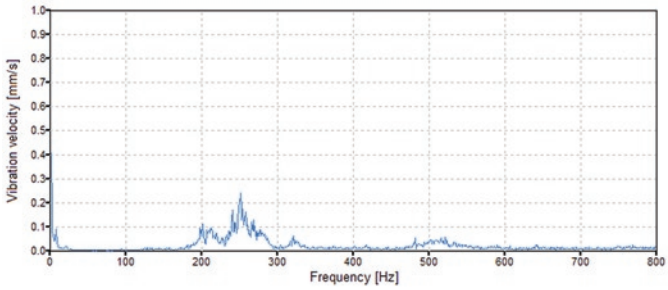


Fig. 10. The vibration velocity spectrum of the microturbine with a damaged thrust bearing, measured on the casing at a rotational speed of 15,060 rpm (251 Hz)

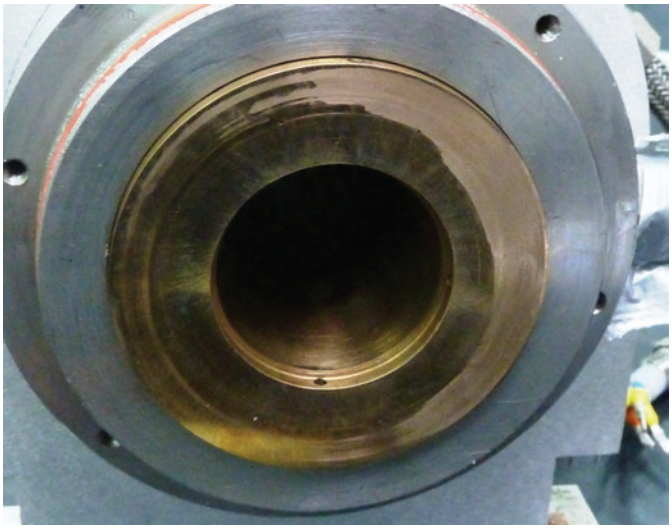


Fig. 11. Gas thrust bearing with a damaged sliding surface

3.4. Research on the microturbine with a bent shaft

This section describes another case of malfunctioning of the 2.5 kW ORC microturbine. It was observed that in addition to the synchronous component (1X), the vibration spectrum also included the harmonic component (2X) with a frequency twice as high as the current rotational frequency of the rotor. The amplitude of the 2X component was generally two times lower than that of the 1X component, regardless of the rotational speed. Two exemplary spectral graphs, showing such a distribution of vibration amplitudes, are in Fig. 12 and Fig. 13. The results demonstrated on these graphs were obtained during measurements at 11,280 rpm and 13,260 rpm. With regard to the second graph (Fig. 13), we can notice that in addition to the 2X component, the vibration spectrum also has a component with a frequency three times higher than the basic frequency. However, the amplitude of the 3X component was much lower than the amplitudes of the 1X and 2X components and, therefore, had a negligible impact on the overall vibration level, on the basis of which the dynamic state of machines is evaluated.

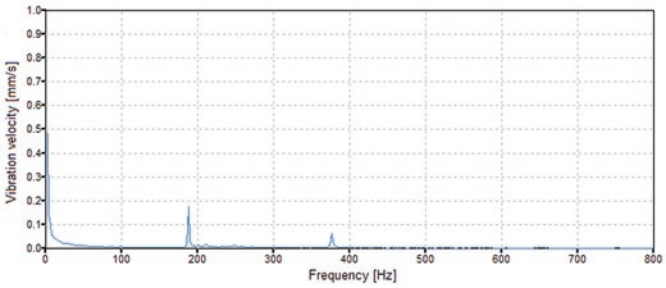


Fig. 12. The vibration velocity spectrum of the microturbine with a bent shaft, measured on the casing at a rotational speed of 11,280 rpm (188 Hz)

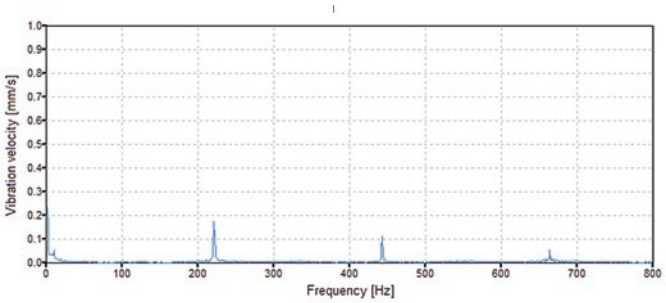


Fig. 13. Vibration velocity spectrum of the microturbine with a bent shaft, measured on the casing at a rotational speed of 13,260 rpm (221 Hz)

When the 2X component is present on a vibration spectrum graph and is present after changing the rotational speed, it usually means that either the shaft is bent or there is a misalignment of two mating shafts. Since the rotating machine discussed herein has only one shaft, it was suspected that the bent shaft was the main reason for the appearance of the 2X component. An analysis of the microturbine design, as well as its start-up and operating conditions, confirmed that the failure of the shaft was very likely. The inlet of the fresh vapour that powers the blade system of the microturbine is located only on one side of the casing (see Fig. 1), which results in the rotor not being uniformly heated. The rotor of this machine starts rotating only after there is an appropriate differential pressure between the vapour inlet and outlet. This means that it cannot be evenly heated during the initial start-up period. A similar situation occurs with radial gas bearings, which are powered through small-diameter holes. If the pressure of the lubricating medium is too low, the bearing journals cannot be lifted. This means that the supply holes in the lower parts of the bearings are blocked by the journals and vapour flows through the remaining holes. Only after the pressure increases, causing the journals to move towards the centres of the bearings, a similar amount of vapour starts flowing through all the supply holes. What aggravates the situation more is that two inlets of hot vapour which power the bearings are located on the same side of the casing. When starting the machine, this may cause thermal deformations as the casing is not evenly heated. Since the bearing sleeves are tightly mounted in the casing, its deformations also cause changes in the geometry of the bearing lubricating gaps. All this clearly shows that the microturbine's rotating system is susceptible to thermal deformations and that they may be the reason behind the appearance of diagnostic symptoms associated with a bent shaft.

A detailed study of this phenomenon showed that additional harmonic components (the clearly visible 2X component and the much lower 3X component) appeared mainly during the quick starts of the

microturbine and then disappeared after a few minutes of operation. When the casing, bearings and the rotor were heated over a longer period of time, this problem practically did not exist anymore. Therefore, in order to avoid a temporary deterioration of the dynamic performance, the settings of the test rig control system were changed so that the start of the microturbine always takes places after the temperature distribution is uniform.

4. Summary and conclusions

The results of vibrodiagnostic studies of the prototypical vapour microturbine with a maximum electric power of 2.5 kW have been discussed. The microturbine was designed for small ORC cogeneration systems, which can be used in single-family houses to produce heat and electricity. Experimental tests of the microturbine were performed under laboratory conditions, on the test rig that made it possible to simulate real operating conditions and apply the target working medium. Since the turbomachine tested had relatively low power and compact dimensions, it was possible to conduct research even after different types of defects of the rotating system had been detected. In the case of large steam turbines, such tests never take place because if there is a suspicion of serious damage to their bearings or rotor, the operation must be interrupted immediately. In addition, the costs of dismantling and repairing a large steam turbine are incommensurable. It was possible to carry out extensive research on the micro-power fluid-flow machine (despite the very high rotational speed – above 20,000 rpm), even in cases where its dynamic characteristics raised concerns. Until now, no research of this type, carried out on energy microturbines, has been presented in the scientific literature.

Based on the research carried out, it can be said that all the detected defects occurred due to improper operation, and in none of the cases were they due to faulty construction, improper fabrication or assembly. All the defects that caused the malfunction of the microturbine were detected using vibrodiagnostics. With regard to specific operating states of the machine, it was as follows:

- During the start-up and the initial period of operation, the microturbine was characterised by a very low vibration level and only the component corresponding to the rotational speed of the rotor (1X) was present on the vibration velocity spectrum. According to the ISO 10816 standard, the dynamic state of the tested machine was very good. The vibration level of the casing was typical for new machines, recently put into service.
- After the pump failure, the oil got mixed with the working medium of the ORC system and then entered the microturbine. As a result, sediments appeared on some parts of the rotor, causing, among other things, a significant unbalance of this component of the microturbine. During diagnostic measurements, this resulted in a considerable increase in the synchronous vibration level (by more than 60%); in other words, the 1X component increased significantly.
- In the case of the defective thrust bearing, in which a few supply holes were clogged, the vibration velocity spectrums became chaotic above certain rotational speeds, which may have been indicative of the occurrence of physical contact between the rotating and non-rotating components of the machine. When the speed increased above 6,000 rpm, higher harmonics appeared and dispersed on nearby frequencies. Then, above 12,000 rpm, an atypical and chaotic vibration distribution was observed near the synchronous component.
- The rotor deflection was the last detected defect of the microturbine. However, this posed a problem only during the first moments after start-up. The analysis showed that the problem was probably related to the uneven heating of the rotor and casing. This problem can be solved by careful planning the heating procedure of the machine before its start-up.

Since all the defects that appeared during operation of the microturbine were detected very quickly, no serious damage occurred. Therefore, the microturbine was restored into operation (without any operating limitations), in a relatively short time and with low additional costs. It should also be noted that a substantial increase in the vibration level occurred only in the case of the unbalanced rotor. Other defects were detected using vibration spectrum analysis. As for the malfunctioning gas bearing and the rotor deflection, the overall vibration level was low. Simple diagnostic methods, based solely on constant monitoring the vibration level, would not have been reliable in this case and the microturbine could suffer serious damage.

The defects and their diagnostic symptoms, detected during the laboratory tests, will be used to create a diagnostic system as it seems to be necessary with regard to the target application of the microturbine. The results of the experimental research presented in this article could be useful for all engineers and researchers involved in the maintenance and diagnostics of various types of energy microturbines, including high-speed ORC microturbines used in small cogeneration systems which have become increasingly popular in recent years.

Acknowledgement

The research discussed herein was conducted within the framework of project No. POIG.01.01.02-00-016/08 and the statutory funds of the IMP PAN in Gdańsk. The authors of the article would like to thank the project manager, Prof. Jan Kiciński, for inspiring us to begin research on cogeneration systems and power microturbines. We also thank Prof. Zbigniew Kozanecki, for his supervision of the microturbine design process, and dr Eugeniusz Ihnatowicz, for his contribution to the construction of the Micro Cogeneration Power Plant Laboratory, which is located in the IMP PAN in Gdańsk.

References

1. Barella S, Bellogini M, Boniardi S, Cincera S. Failure analysis of a steam turbine rotor. Engineering Failure Analysis 2011; 18: 1511-1519, https://doi.org/10.1016/j.engfailanal.2011.05.006.
2. Barsali S, De Marco A, Giglioli R, Ludovici G, Possenti A. Dynamic modelling of biomass power plant using micro gas turbine. Renewable Energy 2015; 80: 806-818, https://doi.org/10.1016/j.renene.2015.02.064.
3. Beith R. (ed.), Small and micro combined heat and power (CHP) systems. Cambridge: Woodhed Publishing Limited, 2011, https://doi.org/10.1533/9780857092755.
4. Czmochowski J, Moczko P, Odyjas P, Pietrusiak D. Tests of rotary machines vibrations in steady and unsteady states on the basis of large diameter centrifugal fans. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(2): 211-216.
5. Dominiczak K, Rządkowski R, Radulski W, Szczepanik R. Online prediction of temperature and stress in steam turbine components using neural network. Journal of Engineering for Gas Turbines and Power 2016; 138: 052606-1, https://doi.org/10.1115/1.4031626.
6. Drescher U, Bruggemann D. Fluid selection for the Organic Rankine Cycle (ORC) in biomass power and heat plants. Applied Thermal Engineering 2007; 27(1): 223-228, https://doi.org/10.1016/j.applthermaleng.2006.04.024.
7. Efimov N, Papin V, Bezuglov R. Determination of rotor surfacing time for the vertical microturbine with axial gas-dynamic bearings. Procedia Engineering 2016; 150: 294-299, https://doi.org/10.1016/j.proeng.2016.07.006.

8. Hong D, Joo D, Woo B, Koo D, Ahn C. Unbalance Response Analysis and Experimental Validation of an Ultra High Speed Motor-Generator for Microturbine Generators Considering Balancing. Sensors 2014; 14: 16117-16127, https://doi.org/10.3390/s140916117.
9. Kaczmarczyk T, Żywica G, Ihnatowicz E. Vibroacoustic diagnostics of a radial microturbine and a scroll expander operating in the organic Rankine cycle installation. Journal of Vibration Engineering 2016; 18(6): 4130-4147, https://doi.org/10.21595/jve.2016.17167.
10. Kataoka T, Kishikawa T, Sakata S, Nakagawa T, Ishiguro J. Remote monitoring and failure diagnosis for a microturbine cogeneration system. ASME Turbo Expo 2007, Montreal (Canada), GT2007-27355, https://doi.org/10.1115/GT2007-27355.
11. Keshtkar H, Alimardani A, Abdi B. Optimization of rotor speed variations in microturbines. Energy Procedia 2011; 12: 789-798, https://doi.org/10.1016/j.egypro.2011.10.105.
12. Kiciński J, Żywica G. Steam microturbines in distributed cogeneration, Cham: Springer, 2014, https://doi.org/10.1007/978-3-319-12018-8.
13. Klonowicz P, Witanowski Ł, Jędrzejewski Ł. A turbine based domestic micro ORC system. Energy Procedia 2017; 129: 923-930, https://doi.org/10.1016/j.egypro.2017.09.112.
14. Kozanecka D, Kozanecki Z, Tkacz E, Łagodziński J. Experimental research of oil-free support systems to predict the high-speed rotor bearing dynamics. International Journal of Dynamics and Control 2015; 3(1): 9-16, https://doi.org/10.1007/s40435-014-0074-9.
15. Kozanecki Z, Łagodziński J. Magnetic thrust bearing for the ORC high - speed microturbine. Solid State Phenomena 2013; 198: 348-353, https://doi.org/10.4028/www.scientific.net/SSP.198.348.
16. Kubitz L, Rządkowski R, Gnesin V, Kolodyazhnaya L. Direct integration method in aeroelastic analysis of compressor and turbine rotor blades. Journal of Vibration Engineering & Technologies 2016; 4(1): 37-42.
17. Liu C, Jiang D, Chen J, Chen J. Torsional vibration and fatigue evaluation in repairing the worn shafting of the steam turbine. Engineering Failure Analysis 2012; 26: 1-11, https://doi.org/10.1016/j.engfailanal.2012.06.001.
18. Margo P, Luck R. Energetic and exergetic analysis of waste heat recovery from a microturbine using organic Rankine cycles. International Journal of Energy Research 2013; 37(8): 888-898, https://doi.org/10.1002/er.2891.
19. Otsu Y, Somaya K, Yoshimoto S. High-speed stability of a rigid rotor supported by aerostatic journal bearings with compound restrictors. Tribology International 2011; 44: 9-17, https://doi.org/10.1016/j.triboint.2010.09.007.
20. Pawlik P. Single-number statistical parameters in the assessment of the technical condition of machines operating under variable load. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 164-169, https://doi.org/10.17531/ein.2019.1.19.
21. Peirs J, Reynaerts D, Verplaetsen F. A microturbine for electric power generation. Sensors and Actuators 2004; 113: 86-93, https://doi.org/10.1016/j.sna.2004.01.003.
22. Poursaeidi E, Mohammadi Arhani M. Failure investigation of an auxiliary steam turbine. Engineering Failure Analysis 2010; 17: 1328-1336, https://doi.org/10.1016/j.engfailanal.2010.03.006.
23. Poursaeidi E, Taheri M, Farhangi A. Non-uniform temperature distribution of turbine casing and its effect on turbine casing distortion. Applied Thermal Engineering 2014, 71: 433-444, https://doi.org/10.1016/j.applthermaleng.2014.07.019.
24. Wang W, Buhl P, Klenk A, Liu Y, The effect of in-service steam temperature transients on the damage behavior of a steam turbine rotor. International Journal of Fatigue 2016; 87: 471-483, https://doi.org/10.1016/j.ijfatigue.2016.02.040.
25. Witek L, Orkisz M, Wygonik P, Musili D, Kowalski T. Fracture analysis of a turbine casing. Engineering Failure Analysis 2011; 18: 914-923, https://doi.org/10.1016/j.engfailanal.2010.11.005.
26. Zhang D, Xie Y, Feng Z. An investigation on dynamic characteristics of a high speed rotor with complex structure for microturbine test rig. ASME Turbo Expo 2008, Berlin (Germany), GT2008-50411, https://doi.org/10.1115/GT2008-50411.
27. Ziegler D, Puccinelli M, Bergallo M, Picasso A. Investigation of turbine blade failure in a thermal power plant. Case Studies in Engineering Failure Analysis 2013; 1: 192-199, https://doi.org/10.1016/j.csefa.2013.07.002.
28. Żywica G, Bagiński P. Investigation of gas foil bearings with an adaptive and non-linear structure. Acta Mechanica et Automatica 2019; 13(1): 5-10, https://doi.org/10.2478/ama-2019-0001.
29. Żywica G, Kaczmarczyk T, Ihnatowicz E, Turzyński T. Experimental investigation of the domestic CHP ORC system in transient operating conditions. Energy Procedia 2017; 129: 637-643, https://doi.org/10.1016/j.egypro.2017.09.123.
30. Żywica G, Kiciński J. The influence of selected design and operating parameters on the dynamics of the steam micro-turbine. Open Engineering 2015; 5: 385-398, https://doi.org/10.1515/eng-2015-0038.

Grzegorz ŻYWICA
Tomasz Z. KACZMARCZYK
Department of Turbine Dynamics and Diagnostics
Institute of Fluid Flow Machinery, Polish Academy of Sciences
Fiszera 14 str., 80-231 Gdansk, Poland

E-mails: gzywica@imp.gda.pl, tkaczmarczyk@imp.gda.pl

Article citation info:
KOZŁOWSKI E, MAZURKIEWICZ D, ŻABIŃSKI T, PRUCNAL S, SĘP J. Assessment model of cutting tool condition for real-time supervision system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 679–685, http://dx.doi.org/10.17531/ein.2019.4.18.

Edward KOZŁOWSKI
Dariusz MAZURKIEWICZ
Tomasz ŻABIŃSKI
Sławomir PRUCNAL
Jarosław SĘP

ASSESSMENT MODEL OF CUTTING TOOL CONDITION FOR REAL-TIME SUPERVISION SYSTEM

MODEL OCENY STANU NARZĘDZIA SKRAWAJĄCEGO DLA SYSTEMU NADZORU W CZASIE RZECZYWISTYM

Further development of manufacturing technology, in particular machining requires the search for new innovative technological solutions. This applies in particular to the advanced processing of measurement data from diagnostic and monitoring systems. The increasing amount of data collected by the embedded measurement systems requires development of effective analytical tools to efficiently transform the data into knowledge and implement autonomous machine tools of the future. This issue is of particular importance to assess the condition of the tool and predict its durability, which are crucial for reliability and quality of the manufacturing process. Therefore, a mathematical model was developed to enable effective, real-time classification of the cutting blade status. The model was verified based on real measurement data from an industrial machine tool.

Keywords: *predictive maintenance, logistic regression, elasticnet, maximum likelihood method, ROC, AUC.*

Dalszy rozwój inżynierii produkcji, w szczególności obróbki skrawaniem, wymaga poszukiwania nowych innowacyjnych rozwiązań technologicznych. Dotyczy to w szczególności zaawansowanego przetwarzania danych pomiarowych pochodzących z systemów diagnostycznych i monitorujących. Rosnąca ilość danych gromadzonych przez wbudowane systemy pomiarowe wymaga opracowania skutecznych narzędzi analitycznych, aby efektywnie przekształcać dane w wiedzę i wdrażać autonomiczne obrabiarki przyszłości. Kwestia ta ma szczególne znaczenie dla oceny stanu narzędzia i przewidywania jego trwałości, które są kluczowe dla niezawodności i jakości procesu produkcyjnego. Dlatego opracowano nowy model matematyczny, którego zadaniem jest skuteczna klasyfikacja stanu ostrza narzędzia skrawającego realizowana w czasie rzeczywistym. Opracowany model został zweryfikowany na podstawie rzeczywistych danych pomiarowych z przemysłowej obrabiarki.

Słowa kluczowe: *predykcyjne utrzymanie ruchu, regresja logistyczna, elasticnet, metoda największej wiarygodności, ROC, AUC.*

1. Introduction

Machining as a manufacturing technology has invariably played a significant role in the manufacturing processes of many enterprises. It is estimated [3] that expenditure on machining account for approximately 5% of the GDP in the developed countries. Therefore, machining technology is constantly evolving. It results from numerous research concerning i.e. the accuracy of the machined parts [5], or the stability of the high-speed machining process [30].

Despite a number of research and innovations regarding technologically advanced cutting tools, or more demanding materials to be machined, further striving to increase productivity and quality decreasing total costs at the same time, requires search for innovative solutions, including those of an optimizing character. Therefore, recently, the number of scientific research in the field of machining is growing. They concern advanced processing of collected measurement data coming from diagnostic and monitoring systems of technological machines. On the one hand, this is the result of the rapid development of measurement and analytical techniques [10, 17, 23], and the growing importance of broadly understood durability and reliability. On the other hand, it is the result of expectations related to the implementation of solutions based on the idea of Industry 4.0. Machine to machine communication, smart technologies, or the need to develop cyber-physical systems (CPS), taking into account the broadly understood principle of sustainable development [14-16], pose a number of new research challenges. According to Lee at al. [22], recent advances in manufacturing industry have paved way for a systematical deployment of CPS systems, within which information from all related perspectives is closely monitored and synchronized between the physical factory floor and the cyber computational space. This requires advanced information analytics for networked machines, which finally will be able to perform more efficiently and collaboratively.

Currently available advanced technological solutions in measurement sensors and data collection and processing systems [24-26, 28, 29] as well as widespread use of industrial computer networks open up an opportunity for potential future smart factories. However, the increasing amount of collected data requires effective analytical tools

[6]. Vast amount of research are also conducted in this area, however, they are mainly theoretical considerations, where new methods or mathematical models are usually verified only based on simulation data. As Arrazola et al. [3] noted that industry application of mathematical models is very limited in manufacturing technology analyzed from the point of view modelling of metal machining operations, due to the fact that direct application of currently available predictive models for specific operations on the shop floor is limited, as most models developed by researchers are only laboratory-validated, and not shop floor-tested. Taking into account the above factors as well as the growing role of reliability and quality of the manufacturing process, in which the durability of the cutting tool plays a key role, a mathematical model was developed to effectively classify the cutting blade. The model was verified based on actual measurement data. This way contribution to industrial data elaboration for predictive maintenance was successfully provided.

2. Problem formulation

The study primarily aimed to determine the cutter state. The answer to the question if the cutter is sharp or not is necessary to define its function, since the response variable is qualitative. Thus the values of this function should be included in a set of two elements describing the possible states of the cutter. The response variable has a categorical value in the case considered. Prediction of cutter state based on data obtained from sensors (accelerometers, microphones, etc.) can be referred to as classification problem.

Numerous classification techniques (classifiers) might be used to predict inclusion in the appropriate class. Logistic regression was used to determine the cutter state. In logistic regression allows to calculate the probability with which the response variable belongs to appropriate category. Therefore, instead of determining cutter state, the probability of each possible state was estimated. In other words, application of logistic regression allows to determine the distribution of response variable based on observation of input variables. Some observable input variables are strongly correlated which will be discussed later in the paper. The elasticnet method was used to minimize this problem. The following subsections present the mathematical aspects required to build a classifier of cutter state.

2.1. Logistic regression

Let us consider the data set, where the realization of the response variable belongs to a binary set. For any finite element we analyze the training set $D=\left\{\left(x_{(i)},y_i\right)\right\}_{1\leqslant i\leqslant n}$, where $\left\{x_{(i)}\right\}_{1\leqslant i\leqslant n}$ denotes a series of input variables, $\left\{y_i\right\}_{1\leqslant i\leqslant n}$ is a series of response variable, where $x_{(i)}\in R^m$, $y_i\in\{0,1\}$ for $1\leqslant i\leqslant n$, denote number of samples, m denotes a number of measurements obtained from transducers (sensors). If the cutter is blunt, then we take $y_i=1$ otherwise we put $y_i=0$. The training set can be presented as $D=\{Y,X\}$, where:

$$Y=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},\quad X=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}=\begin{bmatrix} x_{(1)} \\ x_{(2)} \\ \vdots \\ x_{(n)} \end{bmatrix}.$$

Observing the signal $x_{(i)}\in R^m$ obtained from sensors, it is necessary to classify the cutter state. The task is to find such a classifier $f:R^m\to\{0,1\}$, which would allow to classify the cutter into categories $y=1$ or $y=0$ based on observation $x\in R^m$.

Let (Ω,F,P) be a probability space. On this space a random variable Y with binomial distribution, i.e. $Y:\Omega\to\{0,1\}$ (see e.g. [27]) is defined. In the presented case logistic regression is a regression model, where response variable Y has a binomial distribution. Logistic regression (see e.g. [4, 9, 13]) describes probability of realization of dependent variable Y based on observation of input variables X. Therefore, it is necessary to determine success $P(Y=1|X)$ and defeat $P(Y=0|X)$ probabilities accordingly. In literature [4, 12] the formula:

$$\Theta(X)=\frac{P(Y=1|X)}{P(Y=0|X)}=\frac{P(Y=1|X)}{1-P(Y=1|X)} \tag{1}$$

is called the odds. Thus, the odds is defined as the ratio of success probability to defeat probability. The aim of logistic regression consist in determining the probability of success $p(X)=P(Y=1|X)$ based on observation X. Because the probability of success is $p(X)\in(0,1)$, therefore from formula (1) the odds is $\Theta(X)\in(0,\infty)$ but $ln(\Theta(X))\in(-\infty,\infty)$. The logarithm of odds is called log-odds or logit.

For models of logistic regression the linear dependencies between logit and input variables are analyzed:

$$ln\Theta(X)=ln\left(\frac{p(\beta,X)}{1-p(\beta,X)}\right)=X\beta\ , \tag{2}$$

where $\beta=(\beta_1,\ldots,\beta_m)\in R^m$. When linear system (2) contains an intercept, then in matrix X the column that corresponds to intercept contains ones. From (2) a success probability is calculated as follows:

$$p(\beta,X)=\frac{e^{X\beta}}{1+e^{X\beta}}\ . \tag{3}$$

The maximum likelihood method is usually used to estimate the unknown parameters β in logistic regression (3). Thus the likelihood function is defined as:

$$L(\beta,Y,X)=\prod_{i=1}^{n}p\left(\beta,x_{(i)}\right)^{y_i}\left(1-p\left(\beta,x_{(i)}\right)\right)^{1-y_i}\ . \tag{4}$$

The application of maximum likelihood method consists in solving the task:

$$\max_{\beta} L(\beta,Y,X)\ . \tag{5}$$

As a result the estimators of unknown parameters β for system (2) are obtained. Instead of solving the task (4), the auxiliary task needs to be solved:

$$\max_{\beta} l(\beta,Y,X), \tag{6}$$

where the objective function is defined as the logarithm of likelihood function:

$$l(\beta, Y, X) = \sum_{i=1}^{n} \left(y_i x_{(i)} \beta - ln\left(1 + e^{x_{(i)}\beta}\right)\right). \quad (7)$$

To solve the auxiliary task (6) (determine the unknown parameters β) Newton-Raphson algorithm was applied. Application of this algorithm follows that the unknown parameters β are estimated iteratively. In the step $j+1$ the estimators are determined from the formula:

$$\beta_{j+1} = \beta_j + \left(\frac{\partial^2 l}{\partial\beta\partial\beta^T}(\beta_j)\right)^{-1} \frac{\partial l}{\partial\beta}(\beta_j),$$

where $\frac{\partial l}{\partial\beta}(\beta)$, $\frac{\partial^2 l}{\partial\beta\partial\beta^T}(\beta)$ denote a first and second partial derivatives of the objective function (7).

2.2. Elasticnet

Usually the measurements obtained from sensors are correlated (referred to as the multicollinearity problem). If the input variables (predictors) in linear system (2) are correlated, the direct solution of the task (6) based on the application of Newton-Raphson algorithm does not bring about the expected effect. Additionally, the forecasts based on this model are unstable. Thus the problem depends on the selection of appropriate predictors, which should be included in the regression model (2). On the one hand, these predictors should influence the value of response variable, on the other they should not generate multicollinearity.

In literature there are many techniques (e.g. singular value decomposition, regularization, least angle regression) to solve the problem (6) of multicollinearity. One of possible ways to reduce multicollinearity between predictors is the application of the elasticnet method (see e.g. [12, 13]). This method consists in including the penalty, which depends on values of estimators, in objective function. This technique implies a shrinkage of estimators of unknown parameters. From above when predictors are correlated then we solve the task:

$$\max_{\beta} \sum_{i=1}^{n} \left(y_i x_{(i)} \beta - ln\left(1 + e^{x_{(i)}\beta}\right)\right) - \lambda P_\alpha(\beta) \quad (8)$$

where $\lambda > 0$ and value $P_\alpha(\beta)$ denote the penalty. For $0 \leqslant \alpha \leqslant 1$ the penalty $P_\alpha(\beta)$ is defined as a linear combination of vector norm of estimators β in spaces L_1, L_2 and given by the formula $P_\alpha(\beta) = \frac{1-\alpha}{2}\beta_{L_2} + \alpha\beta_{L_1}$. If $\alpha = 0$, then a classical Tikhonov regularization (ridge regression) is used, while $\alpha = 1$, then Least Absolute Shrinkage and Selection Operator (LASSO). The elasticnet is a connection between ridge regression and LASSO. As will discussed later, the application of elasticnet method allowed to receive classifier based on logistic regression (2) with more accurate and stable detection of cutter state.

2.3. Acoustic signal analysis

Acoustic signal properties were identified by correlation analysis which is related to spectral analysis (see e.g. [11]). Therefore, the time series $\{x_t\}_{t\in N}$ were considered which denote acoustic pressure and is (weakly) stationary in a broad sense with realizations in the set of real numbers R. The autocovariance function of the time series is determined as follows:

$$\gamma_\tau = E(x_t - Ex_t)(x_{t+\tau} - Ex_t) \quad (9)$$

and the autocorrelation function is determined as:

$$r_\tau = \frac{\gamma_\tau}{\gamma_0} \quad (10)$$

for any lag $\tau \in N$. Below two theorems are presented which can be helpful to distinguish acoustic signals.

Theorem 1 (Herglotz). Let γ_τ , $\tau \in N$ denote the autocovariance function of weakly stationary time series. There exists right continuous and non decreasing function $F:[-\pi,\pi] \to [0,\infty)$ such that $F(-\pi) = 0$ and:

$$\gamma_\tau = \int_{-\pi}^{\pi} e^{i\omega\tau} dF(\omega) \quad (11)$$

The proof of Theorem 1 is given in [27]. The function $F(\omega), \omega \in [-\pi,\pi]$ is a spectral function, however if:

$$F(\omega) = \int_{-\pi}^{\omega} f(s)ds \quad (12)$$

the function $f(s)$ is a spectral density function. The relationship between the autocovariance function γ_τ , $\tau \in N$ and the spectral density function $f(s)$ is given below.

Theorem 2. If the autocovariance function γ_τ , $\tau \in N$ has realization in the set of real numbers, then the spectral density function $f(\omega), \omega \in [-\pi,\pi]$ is defined as follows:

$$f(\omega) = \frac{1}{2\pi}\left(\gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j cos(j\omega)\right) = \frac{\gamma_0}{2\pi}\left(1 + 2\sum_{j=1}^{\infty} r_j cos(j\omega)\right) \quad (13)$$

The proof of Theorem 2 can be found e.g. [21, 27]. From theorem 2 for needs to make classifier, each acoustic signal was identified by correlation sequence $\{r_\tau\}_{1\le t\le k}$. The stationary property was checked for each acoustic signal by application Augmented Dickey-Fuller (ADF) test (see e.g. [11, 18]). Additionally, the significance of correlation sequence $\{r_\tau\}_{1\le t\le k}$ for acoustic signal $\{x_t\}_{t\in N}$ was analyzed by application of Ljung-Box test (see e.g. [11]).

3. Numerical example

The aim of research was to analyse the possibility of designing a classifier, which would recognize the cutter state. In order to develop a classifier, 2173 signals (series of acoustic pressure) obtained from microphone were analyzed, where 937 cases were concerned for blunt cutters and 1236 for sharp cutters. Each series $\{x_t^j\}_{1\le t\le n}$ had 16000 measurements ($n = 16000$) collected with 25 kHz sampling frequency and was identified by sequence of correlation values $r^j = \{r_t^j\}_{1\le t\le m} \in [-1,1]^m$ for $0 \le j \le 2173$. It was assumed that the maximal lag is equal to 200 ($m = 200$) [31]. Exemplary realization of signal and sequence of correlation is depicted on Figure 1.

Remark 3. For each analyzed acoustic series $\left\{x_t^j\right\}_{1\le t\le n}$, $0\le j\le 2173$ an ADF test was performed. The results show that the probability of null hypothesis H_0 (the series is non-stationary) does not exceed 0.01 for every series. Hence, we accept that the series describing acoustic signals are stationary.

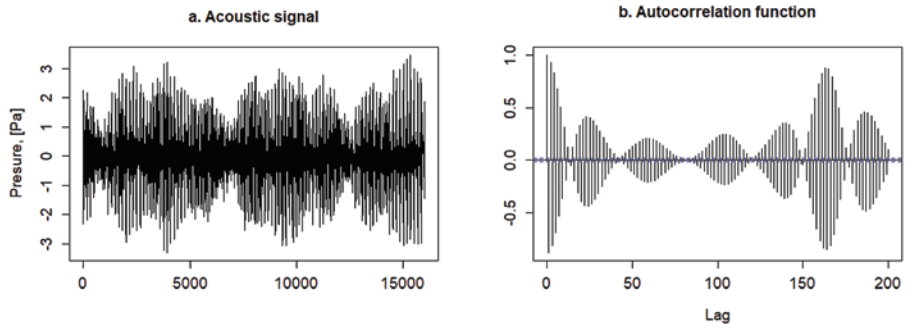


Fig.1 . Realization and autocorrelation for an acoustic signal $\left\{x_t^j\right\}_{1\le t\le n}$

Remark 4. Additionally, Ljung-Box test was performed to examine the null hypothesis H_0, that elements of acoustic series $\left\{x_t^j\right\}_{1\le t\le n}$, $0\le j\le 2173$ are independent. The results show that for each of the analyzed series the probability of null hypothesis H_0 does not exceed 0.01, thus the elements in acoustic series are not independent. From above there exist significant correlations between the elements for lags $\tau\ge 0$)

As a result, learning dataset $D=\left\{\left(r^j,y_j\right): r^j\in[-1,1]^m, y_j\in\{0,1\}, 1\le j\le 2173\right\}$ was created, where $r^j=\left\{r_t^j\right\}_{1\le t\le m}$ is a sequence of values of the autocorrelation function for $j-$th sample, $y_j=0$ for the sharp cutter and $y_j=1$ for the blunt cutter. Thus, observing the sequence , the question if the cutter is blunt must be answered. For this purpose the elasticnet method was applied to estimate unknown parameters in linear model (2) by solving the task (8). Additionally, 10-fold cross validation (see e.g. [12, 13]) was performed to validate stability of the obtained model (thereby to assess of accurately work of model for data).

Below the reconstruction based on application of logistic regression is presented. For correlation sequence of acoustic signal obtained from sensors the probability that the cutter is blunt is calculated as follows:

$$\hat{P}\left(Y=1|r\right)=\frac{e^{r\hat{\beta}}}{1+e^{r\hat{\beta}}} \tag{14}$$

where $\hat{\beta}\in R^m$ denotes the estimator of unknown parameters β for logistic regression (2).

Accordingly, for probability $\hat{P}\left(Y=1|r\right)$, the question of the cutter state must be answered – if it is blunt or sharp, due to the cut-off level (threshold classification) $l\in[0,1]$:

$$state=\begin{cases} sharp, & for\hat{P}\left(Y=1|r\right)<l, \\ blunt, & for\hat{P}\left(Y=1|r\right)\geqslant l \end{cases} \tag{15}$$

Main aim of classification of the cutter state is recognizing that the cutter is blunt. The accuracy (quality of identification) was assessed for different thresholds $0\le l\le 1$. For this purpose the relative error was defined as ratio number of incorrect detection to number of samples (fraction of incorrect detection). The relative error is complement of accuracy (calculated as 1 – accuracy) (see e.g. [1, 2, 7]). Accuracy is a fraction of correct detection from model. Relative error and accuracy give a basic information about goodness of fit of classification model. The smallest relative error corresponds to threshold 0.46. Figure 2 depicts the dependence between relative error and threshold. Below basic terminology and ratios to describe the quality of binary classifier are presented. In this paper, Sharp class is treated as a negative example (N) and Blunt as a positive example (P). To create a confusion matrix we determine the following values: TP (True Positive) denotes number of sample with correct detection for blunt cutter, TN (True Negative) - number of sample with correct detection for sharp cutter, FP (False Positive) - number of sample where sharp cutter was recognized as blunt (number of false alarms), FN (False Negative) - number of sample where blunt cutter was recognized as sharp (number of miss). Basic ratios is calculated as follows:

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN},\ Sensivity=\frac{TP}{TP+FN},\ Specificity=\frac{TN}{TN+FP},$$

$$PositivePredictiveValue=\frac{TP}{TP+FP},\ NegativePredictiveValue=\frac{TN}{TN+FN},$$

$$Prevalence=\frac{TP+FN}{TP+TN+FP+FN},\ DetectionRate=\frac{TP}{TP+TN+FP+FN},$$

$$DetectionPrevalence=\frac{TP+FP}{TP+TN+FP+FN},\ BalancedAccuracy=\frac{Sensivity+Specificity}{2},$$

$$FalseAlarmRate=\frac{FP}{TP+FP}.$$

Table 1 presents a confusion matrix (identification result of cutter state). The relative error of classification is equal to $\frac{(77+106)}{2173}\approx 0.0842$ (below 8.5%), thus accuracy is $1-0.0842=0.9158$. For cut-off level 0.5, the relative error did not exceed 8.6%. Figure 3 presents boxplots (quantiles ¼ and ¾, median and outliers) of probability values obtained by application (8) for cutters.

Table 1. Confusion matrix for classification level 0.46

	Reference: blunt	Reference: sharp
Prediction: blunt	860	77
Prediction: sharp	106	1130

Table 2 presents identification ratios for logistic regression model. The sensitivity (recall or probability of detection) has been calculated as proportion of number of exact (relevant) recognitions for blunt cutters to the number of samples which actually had blunt cutter. On the other hand the specificity has been calculated as proportion of number of exact (relevant) recognitions for sharp cutters to the number of samples which had sharp cutter. The positive predicted value (preci-

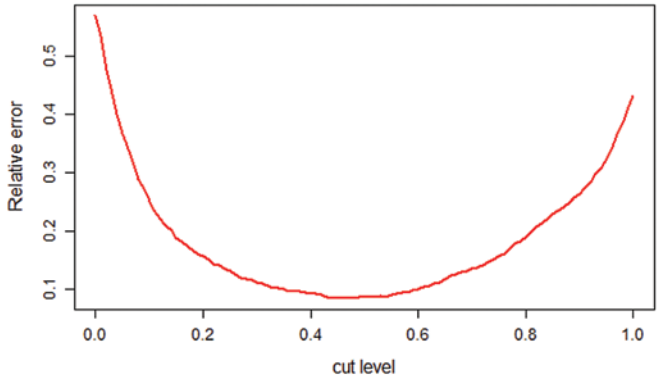


Fig. 2. Relative error dependence on cut level

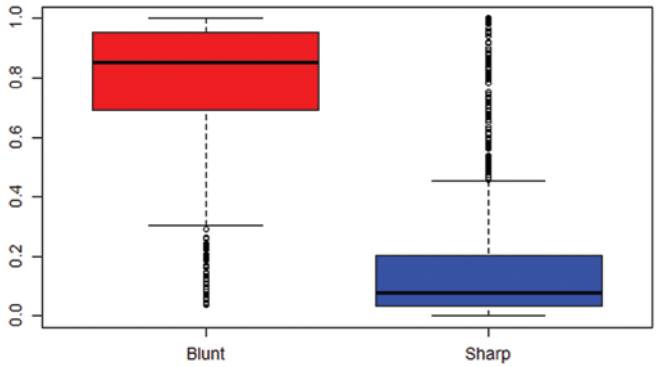


Fig. 3. The result of cutter state detection

sion) is a ratio of number of exact (relevant) recognitions for blunt cutters to the number of samples recognized as blunt. The precision characterizes a purity of performed classifier (purity in retrieval performance). The parameters presented in table 2 are discussed thoroughly in [1, 2, 7, 8]. Additionally, McNemar's chi-squared test was performed for symmetry of rows and columns for confusion matrix presented in table 2. In presented case the McNemar's chi-squared statistic is equal to 4.2842 and p.value is 0.03847. It means that at a significant level 0.05, we have no basis to reject the null hypothesis. Additionally, κ -statistic was calculated. κ -statistic presents the proportion of disagreements expected by chance that did not occur. For recognition based on logistic regression model the parameter κ is equal to 0.829.

Table 2. Basic ratios of quality for logistic regression model

Parameter	Value
Accuracy	0.9158
Sensitivity	0.9178
Specificity	0.9142
Positive Predicted Value	0.8903
Negative Predicted Value	0.9362
Prevalence	0.4312
Detection Rate	0.3958
Detection Prevalence	0.4445
Balanced Accuracy	0.9160
False Alarm Rate	0.0858

Additionally, exact binomial test was also made. Null hypothesis maintains that the probability of success (recognition of blunt cutter) is consistent with prevalence (samples from training data, where the proportion of blunt cutters to all cutters is equal to $\frac{937}{2173} = 0.4312$). At significant level 0.05 we have no basis to reject the null hypothesis.

Figure 3 illustrates the diagnostic ability of classifier based on logistic regression model (see e.g. [7]). Receiver Operating Characteristic (ROC) curve shows the relationship between sensitivity and specificity for every possible cut-off levels. The diagonal line in Figure 3 represents a strategy of randomly guessing cutter state. Thus the presented classifier based on logistic regression is much better then guessing. The bend on the ROC curve corresponds to threshold of 0.465. Area under ROC curve (AUC) denotes general measure of predictiveness. The AUC value for classifier is equal to 0.916.

4. Summary

The extremely dynamic development of technology resulted in wide use of advanced computer systems and sophisticated diagnostic systems equipped with intelligent sensors. Monitoring the parameters of the technological process and diagnostics of key machines and production equipment means that each enterprise collects significant and constantly growing quantities of various types of data. Raw data obtained in this way is a large collection and usually difficult to use directly, requiring appropriate analytical methods. The collected data is, however, a resource of extremely valuable information, which after proper processing and inference should become knowledge, on the basis of which it is possible to increase the efficiency of the company's operation and effectively build a competitive advantage.

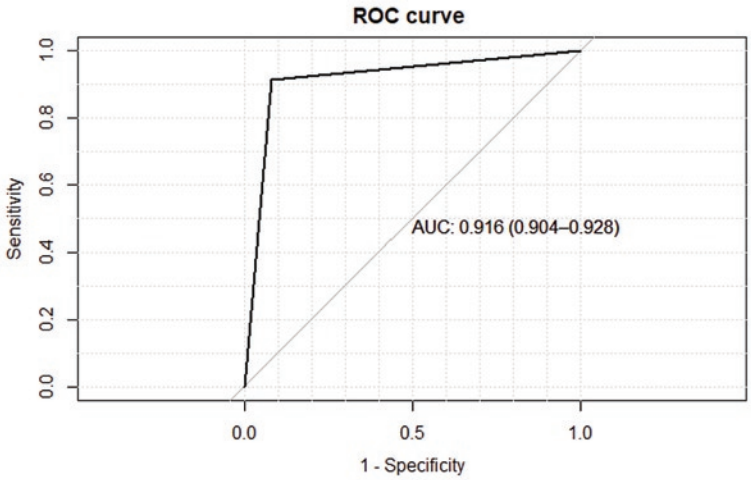


Fig. 4. Receiver operating characteristic curve

Industrial data sets for the implementation of Industry 4.0 solutions are the basis for the functioning of cyber-physical systems, meaning the automation of the processing of collected data. Their functioning requires the development of appropriate analytical tools, especially important for the efficient functioning of CMMS (Computerized Maintenance Management Systems) class predictors. Predictive maintenance is a preventive maintenance method that uses a variety of techniques to inform the owner, service provider or operator about the current and, preferably, future status of their physical resources. The issue of durability and reliability in machining is undoubtedly one of the most important issues also in the context of predictive maintenance and real-time monitoring of technological processes, especially due to the growing role of various diagnostic

and measurement systems, in which every numerically controlled machine tool is equipped. Therefore, the authors undertook the task of developing a solution enabling effective identification of the condition of the cutting tool, allowing further determination of the optimal moment of its replacement. Presented method gives a simple way to detect the cutter state. The study gives promising results - sensitivity and specificity exceed 0.9, furthermore precision is equal to 0.89 and false alarm rate is equal to 0.08. The presented analytical solution, verified on the basis of real data from an industrial machine tool, can be used as part of the system for recognizing the wear rate of the cutting tool during the production process based on the analysis of an acoustic signal or any other symptoms. Obtaining sufficiently high effectiveness of identification, classification and forecasting for the needs of the advanced CMMS class system, however, requires further work. This issue is particularly important for the ongoing work [19, 20] over an innovative prediction tool for failure using mathematical models selected autonomously by an intelligent algorithm based on information criteria, as well as forecast errors and forecast error indicators. As well as the work [31] aimed at developing algorithms for real-time diagnostics and predictions of the state of technological processes.

References

1. Altman D G, Bland J M. Diagnostic tests 1: sensitivity and specificity. British Medical Journal 1994; 308: 1552, https://doi.org/10.1136/bmj.308.6943.1552.
2. Altman D G, Bland J M. Diagnostic tests 2: predictive values. British Medical Journal 1994; 309: 102, https://doi.org/10.1136/bmj.309.6947.102.
3. Arrazola P J, Özel T, Umbrello D, Davies M, Jawahir I S. Recent advances in modelling of metal machining processes. CIRP Annals 2013; 62(2): 695-718, https://doi.org/10.1016/j.cirp.2013.05.006.
4. Balakrishnan N. Handbook of the Logistic Distribution. Marcel Dekker, Inc., 1991, https://doi.org/10.1201/9781482277098.
5. Cheng Q, Sun B, Zhao Y, Gu P. A method to analyze the machining accuracy reliability sensitivity of machine tools based on Fast Markov Chain simulation. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (4): 552-564, http://dx.doi.org/10.17531/ein.2016.4.10.
6. de Jonge B. Maintenance Optimization based on Mathematical Modeling. University of Groningen 2017.
7. Fawcett T. An Introduction to ROC Analysis. Pattern Recognition Letters 2006; 27(8): 861-874, https://doi.org/10.1016/j.patrec.2005.10.010.
8. Fox J, Weisberg S. An R companion to applied regression. SAGE Publications, Inc., 2019.
9. Freedman D A. Statistical Models: Theory And Practice. Cambridge University Press, 2009, https://doi.org/10.1017/CBO9780511815867.
10. Goyal D, Pabla B S. The Vibration Monitoring Methods and Signal Processing Techniques for Structural Health Monitoring: A Review. Archives of Computational Methods in Engineering 2016; 23(4): 585-594, https://doi.org/10.1007/s11831-015-9145-0.
11. Hamilton J. Time Series Analysis. Princeton University Press, 1994.
12. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Springer-Verlag, New York Inc., 2009, https://doi.org/10.1007/978-0-387-84858-7.
13. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning. Springer-Verlag GmbH, 2013, https://doi.org/10.1007/978-1-4614-7138-7.
14. Jasiulewicz-Kaczmarek M, Żywica P. The concept of maintenance sustainability performance assessment by integrating balanced scorecard with non-additive fuzzy integral. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20 (4): 650-661, https://doi.org/10.17531/ein.2018.4.16.
15. Jasiulewicz-Kaczmarek M. Identification of maintenance factors influencing the development of sustainable production processes-a pilot study, IOP Conf. Series: Materials Science and Engineering 2018; 400: 062014, https://doi.org/10.1088/1757-899X/400/6/062014.
16. Jayal A D, Badurdeen F, Dillon O W, Jawahir I S. Sustainable manufacturing: Modeling and optimization challenges at the product, process and system levels. CIRP Journal of Manufacturing Science and Technology 2010; 2(3): 144-152, https://doi.org/10.1016/j.cirpj.2010.03.006.
17. Kant G, Sangwan K S. Prediction and optimization of machining parameters for minimizing power consumption and surface roughness in machining. Journal of Cleaner Production 2014; 83: 151-164. https://doi.org/10.1016/j.jclepro.2014.07.073.
18. Kosicka E, Kozłowski E, Mazurkiewicz D. The use of stationary tests for analysis of monitored residual processes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2015; 17 (4): 604-609, https://doi.org/10.17531/ein.2015.4.17.
19. Kosicka E, Mazurkiewicz D, Gola A. - Multi-criteria decision support in maintenance of machine part. Innowacje w Zarządzaniu i Inżynierii Produkcji, monografia pod red. Ryszarda Knosali, Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, Opole 2016, tom II: 584-593.
20. Kosicka E., Kozłowski E., Mazurkiewicz D. - Intelligent Systems of Forecasting the Failure of Machinery Park and Supporting Fulfilment of Orders of Spare Parts. In book: Intelligent Systems in Production Engineering and Maintenance - ISPEM 2017, Edition: Advances in Intelligent Systems and Computing vol. 637. Publisher: Springer International Publishing, Editors: Anna Burduk, Dariusz Mazurkiewicz, pp. 54-63, https://doi.org/10.1007/978-3-319-64465-3_6.
21. Kozłowski E. Analiza i identyfikacja szeregów czasowych Politechnika Lubelska 2015.
22. Lee J, Bagheri B, Kao H-A. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters 2015; 3: 18-23, https://doi.org/10.1016/j.mfglet.2014.12.001.
23. Leturiondo U, Salgado O, Ciani L, Galar D, Catelani M. Architecture for hybrid modelling and its application to diagnosis and prognosis with missing data. Measurement 2017; 108: 152-162, https://doi.org/10.1016/j.measurement.2017.02.003.
24. Rymarczyk T, Kozłowski E, Kłosowski G. - Electrical impedance tomography in 3D flood embankments testing - elastic net approach. Transactions of the Institute of Measurement and Control 2019; https://doi.org/10.1177/0142331219857374.
25. Rymarczyk T, Nita P, Vejar A, Stefaniak B, Sikora J. - Electrical tomography system for Innovative Imaging and Signal Analysis. Przegląd Elektrotechniczny 2019; 95(6): 133-136, https://doi.org/10.15199/48.2019.06.24.
26. Rymarczyk T., Kłosowski G., Kozłowski E., Tchórzewski P. - Comparison of Selected Machine Learning Algorithms for Industrial Electrical Tomography. Sensors 2019; 9(7):1521, https://doi.org/10.3390/s19071521.
27. Shiryaev A N. Probability-1, 2 . Springer New York, 2016, https://doi.org/10.1007/978-0-387-72206-1.

28. Valis D, Mazurkiewicz D, Forbelska M. Modelling of a Transport Belt Degradation Using State Space Model. In: Proceedings of the 2017 IEEE International Conference on Industrial Engineering & Engineering Management. Singapore: IEEE 2017: 949-953, https://doi.org/10.1109/IEEM.2017.8290032.
29. Valis D, Mazurkiewicz D. Application of selected Levy processes for degradation modelling of long range mine belt using real-time data. Archives of Civil and Mechanical Engineering 2018; 18: 1430-1440, https://doi.org/10.1016/j.acme.2018.05.006.
30. 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, https://doi.org/10.17531/ein.2018.2.19.
31. Żabiński, T., Mączka, T., Kluska, J. Industrial Platform for Rapid Prototyping of Intelligent Diagnostic Systems. In W. Mitkowski, J. Kacprzyk, K. Oprzędkiewicz, P. Skruch (Eds.), Trends in Advanced Intelligent Control, Optimization and Automation. Polish Control Conference, Kraków, Poland 2017: 712-721, https://doi.org/10.1007/978-3-319-60699-6_69.

Edward KOZŁOWSKI
Lublin University of Technology
Faculty of Management
ul. Nadbystrzycka 38, 20-618 Lublin, Poland

Dariusz MAZURKIEWICZ
Lublin University of Technology
Mechanical Engineering Faculty
ul. Nadbystrzycka 38, 20-618 Lublin, Poland

Tomasz ŻABIŃSKI
Sławomir PRUCNAL
Jarosław SĘP
Rzeszów University of Technology
Faculty of Mechanical Engineering and Aeronautics
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland

E-mails: e.kozlovski@pollub.pl, d.mazurkiewicz@pollub.pl, tomz@prz.edu.pl, spktmiop@prz.edu.pl jsztmiop@prz.edu.pl

Article citation info:
GEVORKYAN E, PROKOPIV M, RUCKI M, MOROZOW D. Durability and exploitation performance of cutting tools made out of chromium oxide nanocomposite materials. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 686–691, http://dx.doi.org/10.17531/ein.2019.4.19.

Edvin GEVORKYAN
Mykola PROKOPIV
Mirosław RUCKI
Dmitrij MOROZOW

DURABILITY AND EXPLOITATION PERFORMANCE OF CUTTING TOOLS MADE OUT OF CHROMIUM OXIDE NANOCOMPOSITE MATERIALS

TRWAŁOŚĆ I WŁAŚCIWOŚCI EKSPLOATACYJNE NARZĘDZI SKRAWAJĄCYCH WYKONANYCH Z NANOKOMPOZYTU TLENKU CHROMU

This article is devoted to nanoscale composite materials based on Cr_2O_3 obtained by the activated electric fields sintering procedure. In the paper, exploitative properties of the sintered system of Cr_2O_3 – AlN nanocomposite was examined. Mechanical properties of the material were examined, especially from the perspective of its performance in the cutting tools. In particular, its wear was tested at different cutting speeds, as well as for intermittent hard cutting, and the results were compared with other materials available in the market. Compared to other cutting tools of the same class, Bichromit-R performed the same lifetime for 3-5 times higher cutting speeds, or up to 45% longer lifetime for the same cutting speed. The results lead to the conclusion that composite nanostructure improves substantially exploitation characteristics of the cutting tools.

Keywords: *exploitation, durability, nanocomposite, cutting tool.*

Artykuł jest poświęcony właściwościom eksploatacyjnym materiałów kompozytowych na bazie Cr_2O_3 wytworzonych metodą spiekania w polu elektrycznym. W szczególności poświęcono uwagę nanokompozytowemu spiekowi Cr_2O_3 – AlN wykorzystywanemu do wytwarzania narzędzi skrawających. Zbadano właściwości mechaniczne materiału z uwzględnieniem trwałości ostrzy i powierzchni skrawających. Zbadano zużycie przy różnych prędkościach skrawania w warunkach ciągłych i przerywanych. W porównaniu do ostrzy podobnej klasy, np. Bichromit-R, badane płytki wykazywały podobną trwałość przy wyższych 3 do 5 razy prędkościach skrawania, albo pracowały ok. 45% dłużej przy tych samych prędkościach. Wyniki badań prowadzą do wniosku, że nanostruktura materiału kompozytowego znacząco polepsza właściwości eksploatacyjne ostrzy skrawających.

Słowa kluczowe: *eksploatacja, trwałość, nanokompozyty, narzędzia skrawające.*

1. Introduction

In the engineering systems, even though the lifetime is prolonged, the maintenance cost increases accordingly when fault incurs [7]. In order to reduce expenses, computer-aided maintenance and reliability systems are often applied, as it was reported in case of conveyor belts [19], as well as computer simulation methods [13]. In the context of Industry 4.0, Big Data gains increasing importance [8].

Durability of cutting tools, especially during machining of hard materials, is a subject of many research works [16]. When selecting proper cutting tool for the particular machining task, optimal durability is one of the requirements [1]. One of research directions to prolong cutting tools lifetime is the formation of cutting edge microgeometry which is designed by special processes after grinding or after deposition of the thin layer [28]. There are reports on various layers of nanoscale thickness, e.g. nanocrystalline Al_2O_3 layer deposited by MOCVD on cemented carbide cutting tools [24]. In fact, 85% of all cemented carbide tools are coated [3], but also ceramic materials are coated in order to improve their performance and durability. For instance, Liu et al. proposed novel quaternary coating on the surface of silicon nitride ceramic cutting tool and investigated its dominant wear mechanism [18]. In another reported study, Al_2O_3 was coated on the surface with CaF_2 nanolayer by non-uniform nucleation method, so the mechanical properties of ceramic tools coated with nano-solid lubricant have been significantly improved [4]. It was demonstrated also, that PVD coatings and ALD + PVD hybrid coatings deposited on sialon tool ceramics performed better exploitative properties in comparison with a coating (Ti,Al)N obtained by the conventional method [26].

However, any additional operation of coating, especially with nanolayers, generates increasing costs. Thus, another way to improve durability and performance of ceramic cutting tools is directed to its microstructure formation. It was reported that doping with a small amount of Eu_2O_3 decreases the bulk density and wear resistance of high-alumina ceramics [17]. Since ceramic-matrix composites are outstanding in their ability to withstand high temperatures, in addition their hardness and wear resistance, carbon fiber ceramic-matrix composites are applied, as well as ceramics armed with carbides, nitrides, oxides, and their combinations, including composites with carbon nanotubes and carbon nanofibers [5].

This paper is devoted to the nanocomposite Cr_2O_3 materials produced by the activated electric field sintering procedure. As it will be demonstrated below, its fabrication is cheaper and exploitative properties are better than that of other ceramic cutting tools available in the market.

2. Materials and methods

There are various methods for effective nanopowder consolidation available, and they make possible to obtain materials with a nanosize structure. These methods, such as a hot isostatic pressing (HIP), the high-frequency induction heat sintering (HFIHS), rapid omnidirectional compaction (ROC), pulse plasma sintering (PPS), the ultra high

pressure rapid hot consolidation (UPRC) are quite fully described in works [9, 14, 20, 25].

Each of these methods has some advantages and disadvantages in case of sintering mono and polydispersed electrical conductive and non-conductive nanopowders. Thus, widely applied SPS (Spark Plasma Sintering) method enables to get nanostructured bulk materials from refractory compounds, such as Al_2O_3, SiC, B_4C, $MoSi_2$ etc. [2]. In this method, pulses of current are applied during hot-pressing. In the researches, modified patented field activated sintering method was used with alternating current of 1500-2000 A at voltage 5-10 V [10].

At present, Al_2O_3 is perhaps most widely used material for cutting tools [27]. The chromium oxide (Cr_2O_3) has a crystalline structure similar to Al_2O_3, but it performs slightly higher microhardness 29 GPa compared to Al_2O_3 (28 GPa) because of the strong cohesion.

Chromium oxide nanopowder is obtainable with various methods [22], but there are difficulties in its sintering. In the experiments, the high-density Cr_2O_3 for cutting tools inserts was sintered using typical powders with some additives AlN [15]. This way physico-mechanical properties of materials were considerably improved because of grains nanosize kept by the abovementioned hot-pressing with the electric field [10, 21]. The patented method [12], with reduction of temperature and time of sintering, activates the compression and consolidation mechanisms during sintering process, and also enables to perform compaction of materials otherwise difficult for sintering. As it was demonstrated, short sintering time prevented the growth of grains and ensured improved mechanical properties of the bulk material [11].

Moreover, variation of sintering parameters provides different phase structure of the same material. For example, when the same proportion Cr_2O_3 – 10 wt% AlN was sintered at different temperatures, obtained phase composition of bulk material differed substantially: the sample sintered at T=1500 °C consisted of two phases only, white and grey (marked T1 and T2 in Figure 1), while the one sintered at T=1700 °C had additional dark phase (marked T3 in Figure 1). Table 1 presents the results of quantitative analysis of the obtained phase structures.

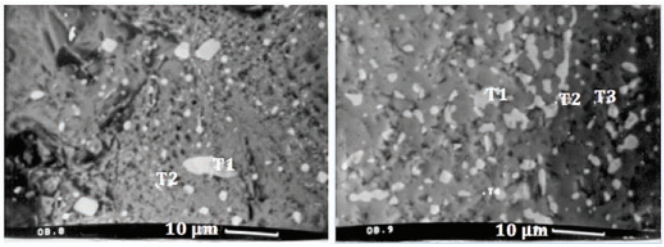


Fig. 1. Photomicrograph of Cr_2O_3 – 10 wt% AlN sintered at T=1500 °C with two phases T1 and T2 (left) and sintered at T=1700 °C with additional phase T3 (right)

Quantitative analysis showed that the dark phase contained large amounts of aluminum, almost two times more that the grey phase. It was found that the dark phase consisted of hard solution $Cr_{1.4}Al_{0.6}O_3$, while the dominant substance in the grey phase was chromium oxide Cr_2O_3.

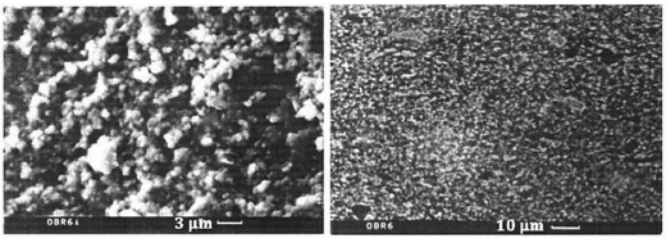


Fig. 2. Fractogram of ceramic fracture Bichromit-R (left) and structure of the surface layer of Bichromit-R after diamond processing (right)

This methodology enabled to obtain the patented material Bichromit-R with nanodispersed structure seen both after fracture test and after diamond grinding, as shown in Figure 2.

Durability tests were carried out during cutting the details made out of steel ШХ-15 (Russian nomenclature), which corresponded with 100Cr6 (ISO standard) and with 52100 (ASTM, USA standard). Hardness of the samples was HRC 58-62. Other steel was used for the evaluation of overall cutting performance of different tool materials. It was steel 30ХГСА (Russian nomenclature), which corresponded with 55 Cr13 (ISO standard) and with 5147 H (ASTM, USA standard) of hardness HRC 58. The machined samples belonged to the group of materials ISO H which contains hardened and tempered steels with hardnesses >45 – 68 HRC. Common steels include carburizing steel (~60 HRC), ball bearing steel (~60 HRC) and tool steel (~68 HRC). Hard types of cast irons include white cast iron (~50 HRC) and ADI/Kymenite (~40 HRC). Constructional steel (40–45 HRC), Mn steel and different types of hardcoatings, i.e. stellite, P/M steel and cemented carbide also belong to this group. Typically, hardness of part machined by turning fall within the range of 55–68 HRC.

No cooling or lubricating was applied. Geometrical features of the sintered inserts and machined samples, as well as the cutting conditions are summarized in the Table 2.

3. Results and discussion

3.1. Mechanical properties

The mechanical properties of the material obtained on the base of Cr_2O_3, called Bichromit-R, were compared with other available ceramic instrumental materials. Since ceramic is a brittle material, increased viscosity is advantageous for its further performance. Figure 3 presents a diagram of stress intensity factors K_{Ic} obtained for different materials typically used for cutting tools inserts manufacturing. Material Bichromit-R performed K_{Ic} above 9 MPa $m^{3/2}$ which indicated higher crack-resistance and hence longer durability than Comp-10, DBC or HC2 materials.

In the Table 3, there are data on main physical characteristics of some cutting tool ceramic materials, compared to Bichromit-R. It is noteworthy that with similar hardness and grain size, Bichromit-R performs better properties than other materials. Above all, its fracture toughness is almost twice higher than for other materials, which indicates high ability of Bichromit-R to resist fractures during cutting

Table 1. Distribution of Cr, Al, O in samples of Cr_2O_3 – 10 wt% AlN sintered at different temperatures

Sintering parameters P=30 MPa	Content of elements, wt%								
	White phase, T1			Grey phase, T2			Dark phase, T3		
	Cr	Al	O	Cr	Al	O	Cr	Al	O
T=1500°C	98.529	0.101	0.292	89.311	6.286	3.906	-	-	-
T=1700°C	96.479	1.729	1.026	81.082	13.172	5.698	71.464	23.735	4.804

Table 2. Geometrical features and cutting conditions in experiments

Geometry of the insert	Intermittent cutting	High-speed cutting	Cutting performance test
Dimensions: 12.5×12.5×4.75 mm l_f = 0.2; r = 0.8 Working angles: γ_0 = -6° ; α_0 = 6° ; φ = 75°; φ_1 = 15°; λ_c = 0°.	f = 0.05 mm/rev; a = 0.1 mm; Cutting speeds from v_c = 60 to 120 m/min	f = 0.075 mm/rev; Cutting speeds from v_c = 25 to 500 m/min	f = 0.5 mm/rev; v_c = 104 m/min

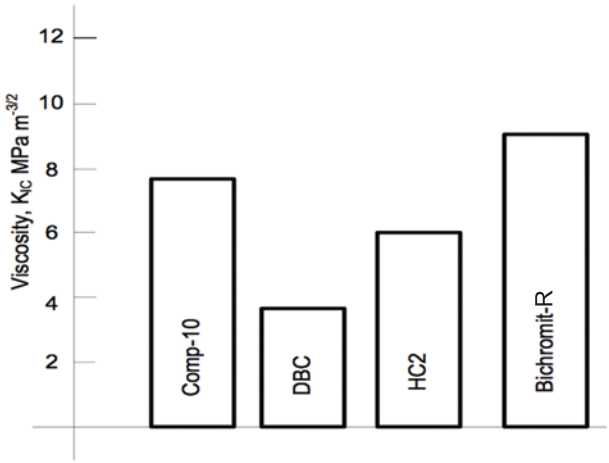


Fig. 3. Fracture toughness diagrams of several cutting tool materials

Table 3. Mechanical characteristics of the Cr_2O_3-based Bichromit-R compared with some ceramic materials

Ceramic type	CC-650 Sweden Al_2O_3	BOK Russia Al_2O_3	Silinite-P Ukraine Si_3N_4	Bichromit-R Ukraine Cr_2O_3
Hardness, HRA	93	92-93	92-94	92-94
Density, g/cm^3	3.97	4.52	3.2-3.4	5.6
Compression strength, MPa	-	-	2500	2600-2800
Bending strength, MPa	480	650	500-700	600-800
Fracture toughness, MPa m$^{1/2}$	6.1	5.6-6.0	4.5	8-10
Grain size, µm	4	2-3	2-3	2-3

operations. This qualifies it for such applications as high speed cutting of hard-tempered cast irons, steel and alloys.

3.2. Durability

The durability comparative tests were performed for intermittent cutting. This type of work conditions is characterized by impact stresses during tool entry, cyclical temperature fluctuation at contact zones between tool and detail, and severe mechanical loading of cutting edge, which usually lead to premature tool failure by fracture [23]. Damage mechanics in intermittent hard cutting can be considered as a combination of microscopic damage and macroscopic fracture of the tool material [6].

The cutting tool made out of Bichromit-R was compared with the one from HC-2 series, based on the aluminum oxide with additions of titanium carbide (Al_2O_3-TiC), produced by NTK. This material is designed and recommended for cutting of hardened steels up to HRC65. In the tests, the steel 5XHM (Russian nomenclature) of HRC 60-63, corresponding with 56CrNiMoV7 (ISO) was machined. In Fig. 4, there are graphs obtained during intermittent cutting at feed f = 0.05 mm/rev; a =0.1 mm.

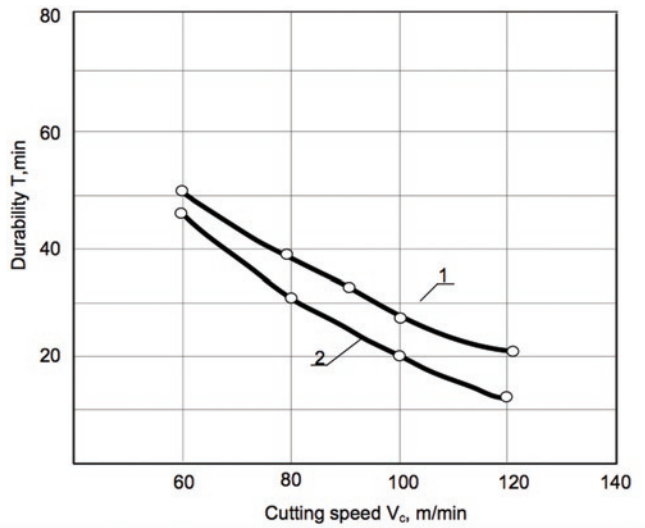


Fig. 4. Durability versus cutting speed during intermittent cutting at feed f = 0.05 mm/rev; a = 0.1 mm; 1 – Bichromit-R, 2 – HC-2

It should be noted that the lifetime of Bichromit-R cutting tools was considerably better than that of HC-2 especially at higher cutting speed. Namely, while at v_c = 60 m/min difference was insufficient, ca. 6%, at doubled speed of 120 m/min Bichromit-R lifetime was ca. 40% longer.

3.3. High-speed cutting

In order to assess the cutting speed influence on the wear of Bichromit-R cutting tools, some tests were carried out. Figure 5 presents the example of results obtained for three different tool materials, namely Bichromit-R, Silinite-P, and BOK-71 (Russian nomenclature). The measure of the wear is the overall path length L [m] of the cut material during machining, before the destruction of the blade. Significantly, the path length ca. L = 20,000 m may be obtained with Silinite-P at cutting speed v_c = 50 m/min, with BOK-71 at v_c = 100 m/min, while with Bichromit-R at v_c = 300 m/min. Moreover, the path length ca. L = 15,000 m may be obtained with Silinite-P at cutting speed v_c = 70 m/min, with BOK-71 at v_c = 130 m/min, while with Bichromit-R even at v_c = 500 m/min. In terms of durability it can be stated that compared with Silinite-P and BOK-71, similar cutting work can be done with Bichromit-R tools, but at the cutting speeds 3-5 times higher.

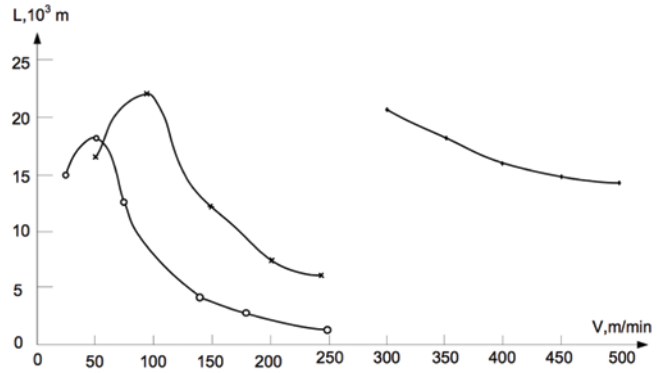


Fig. 5. Cutting speed influence on the wear of a cutting tool h_3 = 0.4mm, during turning of steel ШХ-15 (HRC 58-62) at f = 0.075 mm/rev, and p = 0.2 mm, -o- Silinite-P; -✕- BOK-71; -□- Bichromit-R

3.4. Cutting performance

It should be noted that some operational cutting tests were conducted in-situ by the Volkswagen company (Germany), and they showed that machining with cutting tools made out of Cr_2O_3 material provided high quality of the treated surface of details. That quality was close to the one obtainable by polishing. Other industrial tests were performed at the State Enterprise "Malyshev Plant" (Kharkiv, Ukraine) and they demonstrated that in some turning operations Bichromit-R performed better than other materials available in the Ukrainian market, e.g. "Tomal" cubic boron nitride tools. Thus, ceramics on the basis of chromium oxide could be considered as a new ceramic instrumental material with the high-speed cutting characteristics improved considerably. There are several ways of further improvement of performance of Cr_2O_3-based ceramics, mostly directed to the microstructure features, such as nanoscale grains.

Table 4 presents the comparison of overall performance of different cutting tools in turning operations without cooling at cutting speed v_c = 104 m/min and feed f = 0.5 mm/rev. The machined material was steel 30ХГСА (Russian nomenclature), similar to 4130 (USA) and 25CrMo4 (Germany) of hardness HRC 58, and the materials of cutting tools inserts were typical ceramics of the same class.

The data in Table 4 demonstrates that virtually all tested parameters were better in case of Bichromit-R. Number of passes and total working time was almost twice better, and wear of the tool's back surface was smaller. As a result, roughness of the machined surface was better.

The abovementioned results are mainly attributed to the high fracture toughness discussed in the section 3.1, ensured by the specific sintering technology at smaller temperature and shorter times. It can be assumed that the nanoscale grains of the composite are mainly responsible for the limited crack propagation and unusually high fracture toughness of a ceramic material.

3.5. Physical background

On order to assess the wear of tested cutting tools, the back wear criterion was applied. In case of Bichromit-R, it was h = 0.4 mm. Photomicrographs of the worn tool surface are presented in Figure 6. Like in the surface after the fracture test (Fig 2, left), in the worn surface of the tool submicron structure is clearly seen. Microcracks observable in the micrograph (Fig. 6, right) seem do not develop into large cracks because of nanodispersed structure of material.

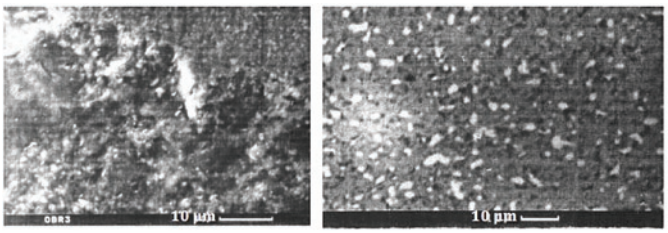


Fig. 6. Microphotographs of the worn surface of the Bichromit-R ceramics (left), and the formation of microcracks at the grain boundaries (right)

Comparative studies of various cutting ceramic materials show that the main reason behind the high wear resistance of Bichromit-R and oxide-carbide ceramics in the processing of steels is the fine-grained structure with submicron elements. Another important feature improving wear resistance is the substructural and dispersed hardening mechanism.

In white Al_2O_3 ceramics, grains do not contain dislocations, which means that grains are not capable to the storage of deformation energy. As a result, micro-destruction of Al_2O_3 grains occurs in the surface layers of the tool. After crack propagation, macroscale wear takes place. This process is slowed down in Bichromit-R due to the features of grain structure.

4. Conclusion

Presented results of the researches demonstrated prolonged durability, higher cutting speeds, smaller wear and better overall performance of cutting tools made out of high-density Cr_2O_3 with some additives AlN, sintered at lower temperatures for a shorter time than usual. Substantial improvement of exploitation characteristics can be attributed to the obtained nanoscale grains inside the bulk material, that are responsible for the increased fracture toughness of a ceramic material, otherwise brittle. Compared to other cutting tools of the same class, Bichromit-R performed the same lifetime for 3-5 times higher cutting speeds, or up to 45% longer lifetime for the same cutting speed.

Acknowledgement

This research work had no specific financial support.

Table 4. Comparative tests of different instrumental materials during machining of the steel 30ХГСА, HRC 58

No.	Cutting insert	Number of passes	Total time	Obtained roughness, Ra	Wear of the tool's back surface, mm	Comment on operation
1	BOK60	11	63	1.25	0.2	Red spiral cutting chip
2	Valenite (USA)	11	63	0.8	0.15	Red spiral cutting chip
3	Hard alloy BK6-OM	5	31.5	2.5	3	squeal, sparking, crumbling
4	Bichromit-R	20	118	0.63	0.1	Red spiral cutting chip after the 15th pass

References

1. Bakša T, Kroupa T, Hanzl P, Zetek M. Durability of Cutting Tools during Machining of Very Hard and Solid Materials. Procedia Engineering 2015; 100: 1414-1423, https://doi.org/10.1016/j.proeng.2015.01.511.
2. Berhard F, Le Gallet S, Spinassou N, Paris S, Gaffet E, Woolman JN, Munir ZA. Dense Nanostructured Materials Obtained by Spark Plasma Sintering and Field Activated Pressure Assisted Synthesis Starting from Mechanically Activated Powder Mixtures. Science of Sintering 2004; 36: 155-164, https://doi.org/10.2298/SOS0403155B.
3. Bobzin K. High-performance coatings for cutting tools. CIRP Journal of Manufacturing Science and Technology 2017; 18: 1-9, https://doi.org/10.1016/j.cirpj.2016.11.004.
4. Chen Zh, Ji L, Guo R, Xu Ch, Li Q. Mechanical properties and microstructure of Al2O3/Ti(C,N)/CaF2@Al2O3 self-lubricating ceramic tool. International Journal of Refractory Metals and Hard Materials 2019; 80: 144-150, https://doi.org/10.1016/j.ijrmhm.2019.01.006.
5. Chung DDL. Carbon Composites: Composites with Carbon Fibers, Nanofibers and Nanotubes. 2nd Edition. Amsterdam: Elsevier, 2017, https://doi.org/10.1016/B978-0-12-804459-9.00001-4.
6. Cui X, Zhao B, Guo J. A review of high-speed intermittent cutting of hardened steel. The International Journal of Advanced Manufacturing Technology 2017; 93(9-12): 3837-3846, https://doi.org/10.1007/s00170-017-0815-y.
7. Dong W, Liu S, Yang X, Wang H, Fang Z. Balancing reliability and maintenance cost rate of multi-state components with fault interval omission. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 37-45, https://doi.org/10.17531/ein.2019.1.5.
8. Fang Y, Tao W, Tee KF. A new computational method for structural reliability with big data. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 159-163, https://doi.org/10.17531/ein.2019.1.18.
9. Fang Zh Z (Ed.). Sintering of Advanced Materials. Oxford: Woodhead Publishing Ltd, 2010, https://doi.org/10.1533/9781845699949.
10. Gevorkian ES, Kodash V Yu. Tungsten carbide cutting tool materials. United States Patent No. 6,617,271 B1 МКИ C 04 B 35/36.
11. Gevorkyan E, Lavrynenko S, Rucki M, Siemiatkowski Z, Kislitsa M. Ceramic cutting tools out of nanostructured refractory compounds. International Journal of Refractory Metals & Hard Materials 2017; 68: 142-144, https://doi.org/10.1016/j.ijrmhm.2017.07.006.
12. Gevorkyan ES, Timofeeva LA, Chishkala VA, Kisly PS. Hot-pressing of tumgsten monocarbide nanopowders with electrical heating. Nanostructural Materials Science 2006; 2: 46-51.
13. Gola A. Reliability analysis of reconfigurable manufacturing system structures using computer simulation methods. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 90-102, https://doi.org/10.17531/ein.2019.1.11.
14. Groza JR, Zavaliangos AK. Sintering activation by external electrical field. Materials Science and Engineering A 2000; 287(2): 171-177, https://doi.org/10.1016/S0921-5093(00)00771-1.
15. Kisly PS, Prokopiv NM, Gevorkyan ES. Raw material for a composite. USSR Invention Certificate No. 759014 V 35/12. 01.05.92.
16. Królczyk G, Gajek M, Legutko S. Predicting the tool life in the dry machining of duplex stainless steel. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(1): 62-65.
17. Liu J, Wu B. Effects of Eu2O3 addition on microstructure, grain-boundary cohesion and wear resistance of high-alumina ceramics. Journal of Alloys and Compounds 2017; 695: 2324-2329, https://doi.org/10.1016/j.jallcom.2016.11.099.
18. Liu W, Li A, Wu H, He R, Huang J, Long Y, Deng X, Wang Q, Wang Ch, Wu Sh. Effects of bias voltage on microstructure, mechanical properties, and wear mechanism of novel quaternary (Ti, Al, Zr)N coating on the surface of silicon nitride ceramic cutting tool. Ceramics International 2016; 42(15): 17693-17697, https://doi.org/10.1016/j.ceramint.2016.08.089.
19. Mazurkiewicz D. Computer-aided maintenance and reliability management systems for conveyor belts. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(3): 377-382.
20. Nersisyan HH, Lee JH, Won CW. SHS for large-scale synthesis method of transition metal nano powders. Int. J. Self Propag. High Temp. Synth. 2003; 12(2): 149-158.
21. Olevsky E A (Ed.). Spark-Plasma Sintering and Related Field-Assisted Powder Consolidation Technologies. Basel: MDPI, 2017.
22. Pei Zh, Zheng X, Li Zh. Progress on Synthesis and Applications of Cr2O3 Nanoparticles. Journal of Nanoscience and Nanotechnology 2016; 16(5): 4655-4671, https://doi.org/10.1166/jnn.2016.12602.
23. Philip PK. Tool wear and tool life in intermittent cutting of hardened steel using conventional hardmetal inserts. International Journal of Machine Tool Design and Research 1978; 18(1): 19-28, https://doi.org/10.1016/0020-7357(78)90016-1.
24. Sawka A, Kwatera A, Woźnicki A, Zasadziński J. Cemented carbide cutting tools life with nanocrystalline Al2O3 layer deposited by MOCVD. Archives of Civil and Mechanical Engineering 2016; 16(3): 351-364, https://doi.org/10.1016/j.acme.2016.01.008.
25. Stanciu LA, Kodash VY, Groza JR. Effects of heating rate on densification and grain growth during field actived sintering of Al2O3 and MoSi2. Metallurgical and Materials Transactions A 2001; 32(10): 2633-2638, https://doi.org/10.1007/s11661-001-0053-6.
26. Staszuk M, Pakuła D, Chladek G, Pawlyta M, Pancielejko M, Czaja P. Investigation of the structure and properties of PVD coatings and ALD + PVD hybrid coatings deposited on sialon tool ceramics. Vacuum 2018; 154: 272-284, https://doi.org/10.1016/j.vacuum.2018.04.032.
27. Wang D, Xue Ch, Cao Y, Zhao J. Fabrication and Cutting Performance of an Al2O3/TiC/TiN Ceramic Cutting Tool in Turning of an Ultra-High-Strength Steel. The International Journal of Advanced Manufacturing Technology 2017; 91(5-8): 1967-1976, https://doi.org/10.1007/s00170-016-9927-z.
28. Zetek M, Česáková I, Švarc V. Increasing Cutting Tool Life when Machining Inconel 718. Procedia Engineering 2014; 69: 1115-1124, https://doi.org/10.1016/j.proeng.2014.03.099.

Edvin GEVORKYAN
Department "Quality, standardization, certification
and manufacturing technology of products"
Ukrainian State University of Railway Transport
Square Feyerbach 7, Kharkov, 61010, Ukraine

Mykola PROKOPIV
Bakul Institute for Superhard Materials,
National Academy of Sciences of Ukraine
2, Avtozavodskaya Str., Kiev, 04074, Ukraine

Mirosław RUCKI
Dmitrij MOROZOW
Faculty of Mechanical Engineering,
Kazimierz Pulaski University of Technology and Humanities in Radom
ul. 54 Stasieckiego, 26-600 Radom, Poland

E-mail: edsgev@gmail.com, keramtex@ukr.net, m.rucki@uthrad.pl,
m.rucki@uthrad.pl

Article citation info:
WALISZYN A, ADAMKIEWICZ A. Studies on resistance to erosion of nickel and its alloys to be used in elements of fluid - flow machines. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (4): 692–696, http://dx.doi.org/10.17531/ein.2019.4.20.

Aleksandr WALISZYN
Andrzej ADAMKIEWICZ

STUDIES ON RESISTANCE TO EROSION OF NICKEL AND ITS ALLOYS TO BE USED IN ELEMENTS OF FLUID - FLOW MACHINES

BADANIA ODPORNOŚCI EROZYJNEJ NIKLU I JEGO STOPU DO ZASTOSOWANIA W ELEMENTACH MASZYN PRZEPŁYWOWYCH*

The article presents results of studies on metal resistance to erosive damage taking place under the influence of hydraulic cavitation. On the basis of earlier research, a hypothesis on fatigue character of erosive wear and a dependence of metal resistance to erosive damage on its crystalline lattice structure has been assumed. To verify this hypothesis, metals with different crystalline lattice structures like steel 45 (flat-centred structure), nickel 200/201 and nickel alloy Monel 400 (hexagonal structure) have been tested at a cavitation-strike stand. Results obtained there confirmed the assumed hypothesis, at the same time justifying the use of nickel protective coatings in fluid-flow machines.

***Keywords**: physics of erosive contamination, resistance, operation, fluid-flow machine, nickel, nickel alloy.*

W artykule przedstawiono wyniki badań odporności metali na uszkodzenia erozyjne zachodzące pod wpływem kawitacji hydraulicznej. Na podstawie wyników wcześniejszych badań, przyjęto hipotezę o zmęczeniowym charakterze zużycia erozyjnego oraz zależności odporności metali na zniszczenia erozyjne od struktury ich sieci krystalicznej. Dla potwierdzenia przyjętej hipotezy na stanowisku kawitacyjno-udarowym sprawdzono metale z różnymi sieciami krystalicznymi: stal 45 (sieć płasko centralna), nikiel 200/201 oraz stop niklu Monel 400 (sieć heksagonalna). Otrzymane wyniki badań potwierdziły przyjętą hipotezę, wskazując tym samym na zasadność stosowania niklowych powłok ochronnych w maszynach przepływowych.

***Słowa kluczowe**: zniszczenia erozyjne, odporność,, maszyna przepływowa, nikiel, stopy niklu.*

1. Physics of cavitation erosion in the cooling systems of diesel engines

Cavitation-erosive damage to fluid-flow machine surfaces washed by fluids and to heat exchange surfaces cooled by fluids is responsible for deterioration of the surface technical condition and decreases their durability. The cause of erosion is usually fluid cavitation in the machine working area. Despite a significant number of studies on cavitation origins of erosive damage to metal surfaces, the physics of fluid flux interaction with the surfaces of protective coating, especially damping ones remains an open case [3, 4, 7].

At present the theory of corrosive – erosive damage to cylinder liners of diesel engines has been widely acclaimed [1, 2]. Practically all researchers believe that the pre-cause of damage to the cooled surfaces of liners and cylinder blocks is the turbulent interaction of liquid with the metal surface being the result of cavitation bubble implosion. Erosive damage to liners and cylinder blocks, which is manifested by creation of clusters of deep cavities, is the result of complex interaction of mechanical and electrochemical processes damaging metal elements i.e. cavitation erosion and electrochemical corrosion. As a result of piston strike when the connecting-rod passes through upper and lower dead centres, cylinder liner experiences vibrations of high frequency which lead to changes in the speed of cooling liquid fluxes on the surfaces of liners and cylinder blocks [1, 13].

Local depressions and increases of liquid pressure which take place at the same time favour severing the continuity of the fluxes and creation of cavitation bubbles in the regions with lowered pressure. The bubbles are filled with vapour, gas or their two-phase mixture.

The change of pressure at any point of the liner surface washed by liquid may be evaluated with a dimensionless coefficient of local discharge [2]:

$$\xi = \left(\frac{\vartheta_0}{\vartheta_i}\right)^2 - 1, \tag{1}$$

where:

ϑ_0 – mean speed of the liquid washing the surface of the liner,
ϑ_i – flux speed in a chosen point i on the surface of the liner.

The largest depression p_i will be in the place where the value of ξ coefficient is maximal:

$$p_i = p_0 - q\xi, \tag{2}$$

where:
p_0 – mean pressure of the liquid,
q – pressure increase speed of the liquid flux:

$$q = \rho\vartheta_i^2 / 2, \tag{3}$$

ρ – liquid density.

The process of cavitation will start when p_i reaches the value equal to the pressure of saturated vapour p_n, at the environment temperature. Generation of cavitation bubbles takes place in the flux washing the surfaces of cylinder liners in the areas where the liquid flows through neckings and has the highest speed, whereas their implosion takes place in the lower speed range where the flux encounters flow

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

resistance or passes through spaces with widening cross-sections. The speed of water flow in cooling systems of diesel ship engines generally does not exceed 2 m/s and does not create conditions favouring hydrodynamic cavitation. Therefore, the basic cause of severing flux continuity should be ascribed to high frequency cylinder liner vibrations [14]. Their presence favours creating cavitation conditions in cooling water systems. Theoretical explanation of this is justified by the fact that when the dynamic pressure ($q = \rho \vartheta_i^2 / 2$) increases, the static pressure of the liquid decreases and conditions favouring flux severing are created. Stretching interactions of the cylinder liner being the result of vibrations add up to this.

Liquid posses volume strength and to stretching stress. At the moment of equilibrium or excessing volume strength under the influence of cylinder liner stretching stress, cavitation processes are initiated – vapour-gas bubbles are being formed. Bubble pulsation frequency is then equal to cylinder liner vibrations.

During the vapour-gas bubble implosion, liquid fluxes of very high speed (up to 34 m/s) appear on its surface, and the pressure of the liquid on the boundary of the imploding bubble and metal surface washed by water may exceed ($5 \cdot 10^{6)}$ Pa. Then the striking energy of the cumulative flux particles is equal to:

$$E = fm\vartheta_k^2 \left[\frac{\eta_1 \omega}{6kT} + \frac{\eta_2}{c^2 \rho} \right]^{-1} \tag{4}$$

where:
m – mass of the liquid of the striking flux,
ϑ – cumulative flux speed,
η_1 , η_2 – kinematic viscosity coefficients of the liquid before and after the bubble implosion,
ω – coefficient of internal friction of liquid molecules,
K – Boltzmann's constant,
T – temperature,
c – sound speed,
f – coefficient accounting for events taking place until the bubble strikes.

The strike of the cumulative liquid flux onto the surface of a washed metal element leads to its plastic deformation and the increase of its hardness – work-hardening. Microcracks being the result of fatigue appear in the strengthened surface metal layers under further striking of cumulative liquid fluxes [8]. Their subsequent development leads to the formation of erosive pits in the form of craters. Figure 1 shows a diagram of metal damage caused by cumulative flux [2].

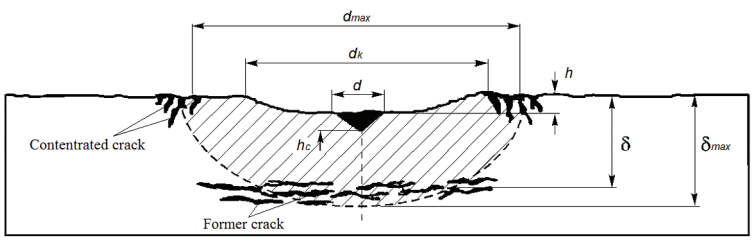


Fig. 1. A diagram of metal damage resulting from erosion cavitation d_{max} – maximum diameter corresponding to the area of concentrated cracks, d_k – pit diameter, d – diameter of oxidation product occurrence area, δ – depth of metal damage

Damage to metal takes place after reaching critical values of stress, which is characteristic for fatigue damage of materials. Frequently, in order to prevent erosive damage, technical literature suggests not very effective methods of increasing the hardness of the surfaces undergoing cavitation erosion. However, in the study [2] cavitation erosion intensity was defined by:

$$J = const \cdot H^n \tag{5}$$

where:
J – erosive wear intensity (mg/mm^{2}hour) of the metal;
H – metal surface hardness (HB);
n – exponent with values ranging from 2,78 (for carbon steels) to 0 (for chromium alloy steels).

The influence of initial steel hardness on the intensity of cavitation erosion is shown in Fig. 2 [12, 13]. Erosive damage to alloy steel 03HG10-10 with lower hardness in comparison to those used in the experiment was much smaller. Thus, in order to increase cavitation erosion resistance, the surface has to be more plastic.

The increase of resistance to cavitation erosion of metal surfaces can be reached throughout:
- making the working flux more laminar;
- covering metal surfaces with protective coatings
- damping vibrations of the element in the case of vibration cavitation.

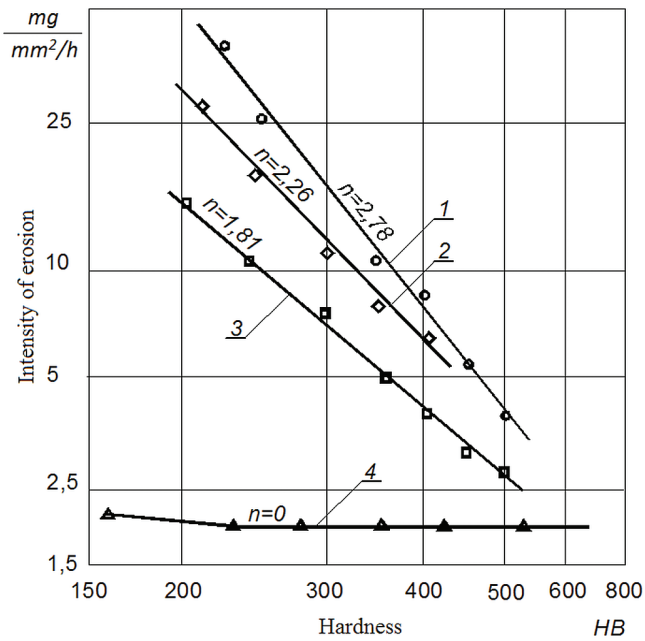


Fig. 2. Intensity of cavitation erosion in relation to initial steel hardness 1– S40 i S1Cr40; Cr5V3; 2 – 1Cr13, 2Cr13, Mn20Si1; 3 – 1Cr18Ni3Mn4Cu2; 4 – 30Cr10Mn10

Laminar flow of liquid fluxes is connected with organising hydraulic systems in which liquid flux speeds do not exceed critical values corresponding with limiting values of the Reynold's criterion. The critical value, marking the beginning of cavitation, is defined by a mathematical relation in the form of (6) [2]:

$$K_k = \frac{K_k^0}{C_j^0} \left(C_j - \Delta C_j^* \right) \tag{6}$$

where: K_k^0 , C_j^0 – cavitation number and resistance coefficient of the by-wall non-laminar layer of the liquid flux:

ΔC_j^* – depending on the structure of the liquid flux, it is defined by:

$$\Delta C_j^* = \frac{1{,}33 \left(1 - \frac{1}{e^F} \right)}{\sqrt{R}} \tag{7}$$

where: e^F – limiting plasticity of metal surface washed by the liquid [2].

An analysis of equations (6-7) for the C_j^* resistance coefficient indicates that when the criterial Reynold's number decreases and when the plasticity of the washed surface increases, the resistance of the outer layer increases. Respectively, cavitation number K_k, which characterises conditions for cavitation phenomena initiation and damage connected with them, decreases.

Taking into account that at the initial moment, work hardening appears in the metal surface layer as a result of the influence of cumulative fluxes which in cyclic strike conditions gets damaged and cracks. Generated microcracks are the seeds of deep pits at then occurring crevice electrochemical corrosion. As a consequence erosive-corrosive damage can be treated as a fatigue-corrosive process. At the same time it can be stated that metal resistance to erosive-corrosive damage to a significant extent depends on the properties of the metal, its chemical composition and the properties of the surrounding medium.

Considering the fatigue character of erosive metal damage which takes place as a result of strikes of the liquid flux at the moment of explosions of vapour-gas cavitation bubbles, the following conclusions can be drawn. Strengthening of the metal surface layer and the resulting work-hardening as can be seen in Figure 1 is the initial stage of erosive damage. In the further development of the process, as a result of appearing fatigue stress, the strengthened metal surface gets cracked and later erosive cavities appear when simultaneously fatigue stress and crevice corrosion occur. Basing on the above conclusions, the most effective method preventing erosive damage to metal surfaces is to be found in the rational choice of mechanical properties of metals used in hydraulic installations. Obviously, the main condition of their application should be their inability to form strengthened surface layers which means that they should possess high plasticity.

2. Choice of study object

Considering maintenance of elasticity of single crystals of structural metals, one should take into account the anisotropy of elastic crystalline modules with the view to determining the differences in elastic deformation when the load is placed at different crystallographic directions. Most metals crystallise in three types of lattice: body-centred, face-centred or hexagonal lattice. These lattices are shown in Figure 3. In all of them, crystal face indexes are marked in the following way: along the X axis as [100]; Y axis - [010] and Z axis - [001]. Indexes of face diagonals are marked as [110] on the X-Y plane, [011] on the Y-Z plane and [101] on the X-Z plane; the body diagonal as [111]. Cast iron, steel and copper belong to metals with the face-centred lattice; nickel, zinc, aluminium and cadmium have hexagonal lattice. An important characteristics of metals from the point of view of their erosive resistance is their plasticity, which is shown in formula (7). Study [2, 3] notes that plasticity of metals with hexagonal lattice is higher than those with face-centred lattice. Deformation of these metals appears as a slide along the [001] plane, which does not lead to strengthening of the surface layer, that is to work-hardening. Deformation of metals with face-centred lattice in [111] and [112] planes is shown in Figure 4.

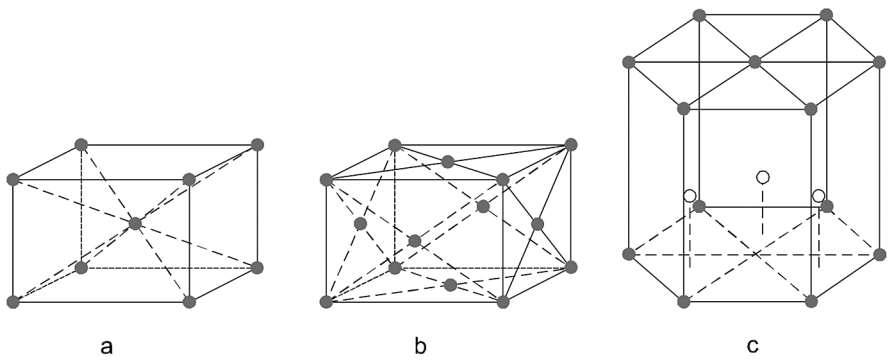


Fig. 3. Elementary cells of close packed lattices: a – body centred lattice; b – face-centred lattice; c – hexagonal lattice

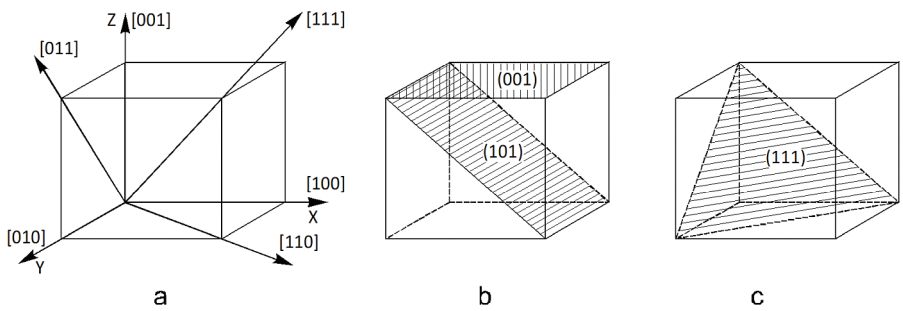
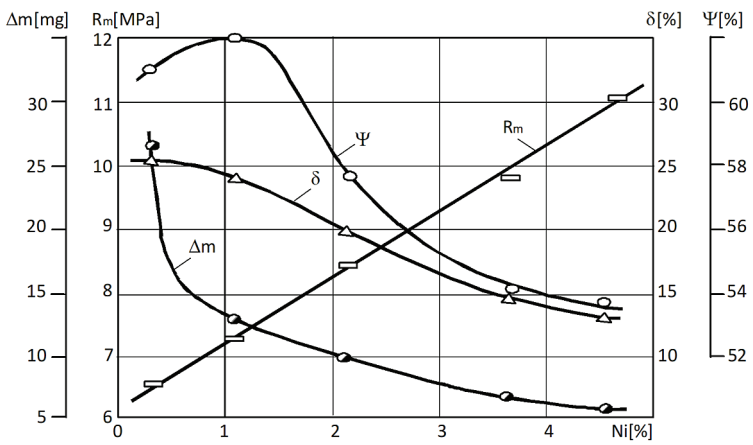


Fig. 4. Crystallographic indexes of deformation directions (a) and shear planes (b, c)

3. Comparative studies on resistance of metals with different crystalline lattice structure and their results

On the basis of an analysis of methods applied for studying metal resistance to cavitation erosive damage [14, 15], the method of ventilated cavitation has been chosen. As cavitation erosion is the effect of crashes of micro-fluxes of liquid after vapour-gas bubbles implosion, the experiment was carried out at a vibration stand using tap water as the working medium.



Δm – mass loss; R_m – temporary resistance; δ – relative elongation; Ψ – relative narrowing

Fig. 5. The influence of nickel content on the properties of alloy steels

Table 1. Chemical composition of samples with nickel 200/201

	C	Si	Mn	S	Co	Cu	Fe	Mg	Ti	Ni	Ni-Co
	0.01	0.04	0.10	<0.01	<0.01	<0.01	<0.01	0.102	0.04	99.67	99.687
Max	0.02	0.15	0.35	0.005	1.0	0.15	0.25	0.15	0.10		
Min											99.6

Table 2. Chemical composition of samples out of monel 400

	C	Si	Mn	Sr	Al	Co	Cu	Fe	Mg	Ti	Ni	Ni-Co
	0.13	0.23	0.94	0.03	<0.01	0.04	32.6	2.06		0.02	63.9	64.007
Max	0.15	0.5	2.0	0.02	0.5	1.0	34.0	2.5		0.3		
Min							28.0	1.0			36.0	63.0

Table 3. Results of studies of metal samples at a vibration stand

Mass (g)	Nickel			Monel			Steel C45		
	Sample 1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3
Mass for studies	9.0722	9.1200	8.8327	8.3740	9.7898	9.4627	8.7153	7.9301	7.8240
Mass after 10-min. exposition	9.0722	9.1200	8.8327	8.3739	9.7897	9.4626	8.7150	7.9298	7.8238
Mass after 30-min. exposition	9.072	9.1200	8.8327	8.3738	9.7896	9.4624	8.7147	7.9296	7.8235
Mass after 60-min. exposition	9.0717	9.1199	8.8325	8.3736	9.7894	9.4623	8.7140	7.9292	7.8231
Hardness of samples for studies (Vickers' method)	95.8	98.3	112.3	125.0	123.4	125.5	155.6	138.5	126.0
Mass decrease of samples during studies	0.0005	0.0001	0.0002	0.0004	0.0003	0.0013	0.0013	0.0009	0.0009
Mean metal decrease	0.00027			0.0007			0.001		
Speed of wear after a 10-min exposition (g/hour)	0			0.0018			0.0016		
Speed of wear after a 30-min exposition (g/hour)	0			0.0014			0.00106		
Speed of wear after a 60-min exposition (g/hour)	0.00027			0.0007			0.001		

Verification of the hypothesis of the fatigue character of erosive damage has been carried out for samples of three metals with different crystalline lattice structure: nickel 200/201, nickel alloy – monel 400 and constructional steel C45 [5, 6, 7]. Chemical composition of the studied nickel 200/201 and monel 400 samples is given in Tables 1 and 2.

Surfaces of samples of studied metals were polished to the roughness of 0,63 μm, after which the hardness of their upper layers was determined applying Vickers' method according to the PN-EN ISO 6507-1:2000 standard. The results, calculated as means of six measurements, are shown in Table 3.

Studies were carried out in compliance with the ASTM G32 standard in the option with an immobile sample in three series: with a 10-minute flux interaction on the sample, 30-minute and 60-minute long interactions. After each series, the samples were weighed using an analytical balance. Measurements of mass decrease enabled determination of the speed of erosive wear of the metals used in the experiment. Results of the studies are listed in Table 3 and graphically presented in Fig 6.

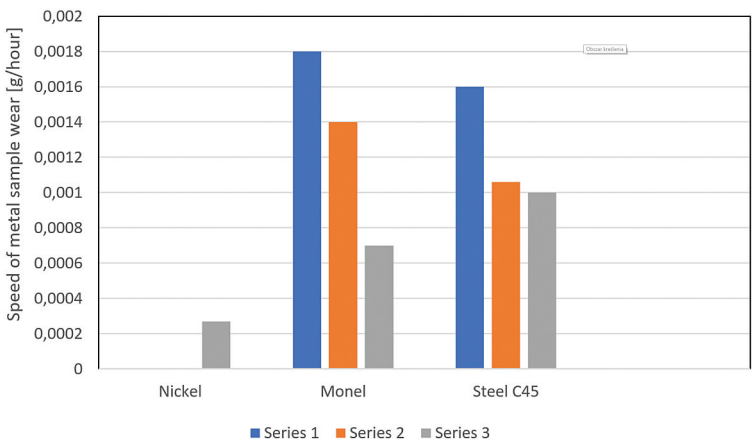


Fig. 6. Speed of erosive wear of metals in the experiment, where series 1 refers to a 10 minute long exposition, series 2 – 30 minute long and series 3 – 60 minute long exposition

4. Summary

Comparing the results of studies, it can be seen that Steel C45 has the lowest resistance to erosive wear in which case the speed of wear is thirty-seven times higher than that of nickel. Monel 400 reaches the values in-between. This fact also supports the assumed hypothesis of effectiveness of using plastic metals to prevent erosive damage in water installations which are prone to cavitation. The second fact supporting this hypothesis is the decrease of steel C45 and monel 400 wear with the increase of experiment time (Fig 6). It is the result of the increase of hardness of the sample upper layers and formation of work-hardening.

Summing up the carried out research, it can be said that prevention and slowing down erosive wear of working surfaces of fluid-flow machine parts may be achieved by using plastic metals to manufacture them.

References

1. Adamkiewicz A, Waliszyn A. Discussion and studies of the properties of a cooling water additive preventing erosive wear of cooled surfaces of ship diesel engines. Eksploatacja i Niezawodność - Maintenance and Reliability 2014; 10(1): 565-570.
2. Adamkiewicz A, Waliszyn A. Studies of erosion resistance of protective coats on the surfaces of machine elements washed with fluids. Advances in Materials Science 2018; 6: 69-76, https://doi.org/10.1515/adms-2017-0033.
3. Amann T, Waidele M, Kailer A. Analysis of mechanical and chemical mechanisms on cavitation erosioncorrosion of steels in salt water using electrochemical methods. Tribology International 2018; 124: 238-246, https://doi.org/10.1016/j.triboint.2018.04.012.
4. Bolewski Ł, Szkodo M, Kmieć M. Cavitation erosion degradation of Belzona® coatings. Advances in Materials Science 2017; 17(1): 22-33, https://doi.org/10.1515/adms-2017-0002.
5. Ciubotariu C R, Secosan E, Marginean G, Frunzaverde D, Campian V C. Experimental Study Regarding the Cavitation and Corrosion Resistance of Stellite 6 and Self-Fluxing Remelted Coatings. Strojniski Vestnik -Journal of Mechanical Engineering 2016; 62 (3): 154-162, https://doi.org/10.5545/sv-jme.2015.2663.
6. Heathcock C J, Protheroe B E, Ball A. Cavitation erosion of stainless steels. Wear 1982; 81(2): 311-327, https://doi.org/10.1016/0043-1648(82)90278-2.
7. Kim J H, Lee M H. A Study on Cavitation Erosion and Corrosion Behavior of Al-, Zn-, Cu-, and Fe-Based Coatings Prepared by Arc Spraying. Journal of Thermal Spray Technology 2010; 19(6): 1224-1230, https://doi.org/10.1007/s11666-010-9521-0.
8. Krella A. Cavitation degradation model of hard thin PVD coatings. Advances in Materials Science 2010; 10(3): 27-36, https://doi.org/10.2478/v10077-010-0010-4.
9. Kumar H, Chittosiya C, Shukla V.N. HVOF Sprayed WC Based Cermet Coating for Mitigation of Cavitation, Erosion & Abrasion in Hydro Turbine Blade, MATERIALS TODAY-PROCEEDINGS 2018; 5(2): 6413-6420. https://doi.org/10.1016/j.matpr.2017.12.253
10. Krumenacker L, Fortes-Patella R, Archer A. Numerical estimation of cavitation intensity. IOP Conference Series-Earth and Environmental Science 2014; 22: Article Number: UNSP 052014, https://doi.org/10.1088/1755-1315/22/5/052014.
11. Kwok C T, Man H C, Cheng F T. Cavitation erosion and damage mechanisms of alloys with duplex structures. Materials Science and Engineering A242 1998: 108-120, https://doi.org/10.1016/S0921-5093(97)00514-5.
12. Steller J, Krella A, Koronowicz J, Janicki W. Towards quantitative assessment of material resistance to cavitation erosion. Wear 2005; 258: 604-613, https://doi.org/10.1016/j.wear.2004.02.015.
13. Waliszyn A, Adamkiewicz A. A metod of vibration damping for diesel engeene cylinder lines to prevent the consequences of erosion. Eksploatacja I Niezawodnosc - Maintenance and Reliability 2018; 20: 371-377, https://doi.org/10.17531/ein.2018.3.4.
14. Yang D, Yu A, Ji B, Zhou J, Luo X. Numerical analyses of ventilated cavitation over a 2-D NACA0015 hydrofoil using two turbulence modeling methods. Journal of Hydrodynamics 2018; 30(2): 345-356, https://doi.org/10.1007/s42241-018-0032-7.
15. Yu A, Luo X, Ji B. Analysis of ventilated cavitation around a cylinder vehicle with nature cavitation using a new simulation method. Science Bulletin 2015; 60(21): 1833-1839, https://doi.org/10.1007/s11434-015-0916-7.

Aleksandr WALISZYN
Andrzej ADAMKIEWICZ
Maritime University of Szczecin, Faculty of Marine Engineering
Wały Chrobrego 1-2, 70-500 Szczecin, Poland

E-mail: a.valishin@am.szczecin.pl, a.adamkiewicz@am.szczecin.pl

The Journal is indexed and abstracted in the Journal Citation Reports (JCR Science Edition), Scopus, Science Citation Index Expanded (SciSearch®) and Index Copernicus International.

The Quarterly appears on the list of journals credited with a high impact factor by the Polish Ministry of Science and Higher Education and is indexed in the Polish Technical Journal Contents database – BAZTECH and the database of the Digital Library Federation.

Ministry of Science and Higher Education
Republic of Poland

Task *„Implementation of procedures ensuring the originality of scientific papers published in the quarterly „Eksploatacja i Niezawodność – Maintenance and Reliability"* financed under contract 532/P-DUN/2018 from the funds of the Minister of Science and Higher Education for science dissemination activities.

All the scientific articles have received two positive reviews from independent reviewers.
Our 2018 Impact Factor is 1.806

Editorial staff:	Dariusz Mazurkiewicz, PhD, DSc (Eng), Associate Professor (Editor-in-Chief, Secretary of the Editorial Board) Tomasz Klepka, PhD, DSc (Eng), Associate Professor (Deputy Editor-in-Chief) Teresa Błachnio-Krolopp, MSc (Eng) (Editorial secretary) Andrzej Koma (Typesetting and text makeup) Krzysztof Olszewski, PhD (Eng) (Webmaster)
Publisher:	Polish Maintenance Society, Warsaw
Scientific patronage:	Polish Academy of Sciences Branch in Lublin
Address for correspondence:	"Eksploatacja i Niezawodność" – Editorial Office ul. Nadbystrzycka 36, 20-618 Lublin, Poland e-mail: office@ein.org.pl http://www.ein.org.pl/
Circulation:	550 copies

INFORMATION FOR AUTHORS

Eksploatacja i Niezawodnosc – Maintenance and Reliability – the journal of the Polish Maintenance Society, under the scientific supervision of the Polish Academy of Sciences (Branch in Lublin), published four times a year.

The scope of the Quarterly

The quarterly *Eksploatacja i Niezawodnosc – Maintenance and Reliability* publishes articles containing original results of experimental research on the durability and reliability of technical objects. We also accept papers presenting theoretical analyses supported by physical interpretation of causes or ones that have been verified empirically. *Eksploatacja i Niezawodność – Maintenance and Reliability* also publishes articles on innovative modeling approaches and research methods regarding the durability and reliability of objects.

The following research areas are particularly relevant to the journal:

1. degradation processes of mechanical and biomechanical systems,
2. diagnosis and prognosis of operational malfunctions and failures.
3. analysis of failure risk/wear,
4. reliability-and-environmental-safety engineering in the design, manufacturing and maintenance of objects,
5. management and rationalization of object maintenance,
6. risk management in the processes of operation and maintenance,
7. the human factor and human reliability in operation and maintenance systems.

Terms and Conditions of Publication

The quarterly *Eksploatacja i Niezawodnosc – Maintenance and Reliability* publishes only original papers written in English or in Polish with an English translation. Translation into English is done by the Authors after they have received information from the Editorial Office about the outcome of the review process and have introduced the necessary modifications in accordance with the suggestions of the referees! Acceptance of papers for publication is based on two independent reviews commissioned by the Editor. The publication fee for one text is 320 EUR (all graphics in color included).

The quarterly *Eksploatacja i Niezawodnosc – Maintenance and Reliability* proceeds entirely online at submission.ein.org.pl

Technical requirements

- After receiving positive reviews and after acceptance of the paper for publication, the text must be submitted in a Microsoft Word document format.
- Drawings and photos should be additionally submitted in the form of high resolution separate graphical files in the TIFF, SVG, AI or JPG formats.
- A manuscript should include: names of authors, title, abstract, and key words that should complement the title and abstract (in Polish and in English), the text in Polish and in English with a clear division into sections (please, do not divide words in the text); tables, drawings, graphs, and photos included in the text should have descriptive two-language captions, if this can be avoided, no formulae and symbols should be inserted into text paragraphs by means of a formula editor; references (written in accordance with the required reference format); author data – first names and surnames along with scientific titles, affiliation, address, phone number, fax, and e-mail address.

The Editor reserves the right to abridge and adjust the manuscripts. All submissions should be accompanied by a submission form.

Detailed instructions to Authors, including evaluation criteria can be found on the journal's website: www.ein.org.pl

Editor contact info

Editorial Office of „Eksploatacja i Niezawodnosc – Maintenance and Reliability"
Nadbystrzycka 36, 20-618 Lublin, Poland
e-mail: office@ein.org.pl

INFORMATION FOR SUBSCRIBERS

Fees

Yearly subscription fee (four issues) is 100 złoty and includes delivery costs. Subscribers receive any additional special issues published during their year of subscription free of charge.

Orders

Subscription orders along with authorization to issue a VAT invoice without receiver's signature should be sent to the Editor's address.

Note

In accordance with the requirements of citation databases, proper citation of publications appearing in our Quarterly should include the full name of the journal in Polish and English without Polish diacritical marks, i.e.,

Eksploatacja i Niezawodnosc – Maintenance and Reliability.

Dr hab. inż. prof. WAT Jacek Paś

Wydział Elektroniki
Wojskowa Akademia Techniczna
gen. S. Kaliskiego 2, 00-908 Warszawa, Polska
E-mail: jacek.pas@wat.edu.pl

Mł. bryg. dr inż. Tomasz Klimczak

Szkoła Główna Służby Pożarniczej
Juliusza Słowackiego 52/54, 01-629 Warszawa, Polska
E-mail: tklimczak@sgsp.edu.pl

Wybrane zagadnienia oceny niezawodnościowo-eksploatacyjnej systemów sygnalizacji pożaru

Słowa kluczowe: *eksploatacja, niezawodność, systemy sygnalizacji pożaru*

Streszczenie: W artykule przeprowadzono analizę problemów eksploatacyjnych i niezawodnościowych, która dotyczy wybranych systemów sygnalizacji pożaru (SSP) o różnej strukturze funkcjonalnej. Systemy te są użytkowane na rozległym obszarze transportowym, w określonym środowisku. Można wyróżnić trzy podstawowe struktury tych systemów - skupiona, rozproszona i mieszana. Dany rodzaj struktury funkcjonalnej systemu, który jest użytkowany w obiekcie (na danym obszarze) jest funkcją zależną od konfiguracji, wewnętrznych połączeń elementów i urządzeń oraz opracowanego scenariusza postępowania na wypadek pożaru. Zastosowanie danej struktury systemu do ochrony pożarowej zależy także od przepisów prawnych warunkujących dopuszczenie danego obiektu (obszaru) do użytkowania. Proces realizacji scenariusza w czasie pożaru jest gwarantowany przez algorytm zaimplementowany w centrali alarmowej oraz innych elementach systemu. Realizacja wszystkich wymagań wobec systemu określonych w danym algorytmie postępowania uwarunkowana jest np. odpowiednią strukturą niezawodnościową i warunkami środowiskowymi. W artykule przedstawiono analizę procesu eksploatacji wybranych SSP, które są użytkowane na obszarze transportowym. Zaprezentowano rzeczywiste wyniki badań procesu eksploatacji, np. czasy trwania naprawy oraz uszkodzenia. Następnie opracowano graf relacji eksploatacyjnych z uwzględnieniem przeprowadzonych badań eksploatacyjnych. Umożliwiło to wyznaczenie zależności pozwalających na określenie parametrów eksploatacyjnych i niezawodnościowych przebywania SSP w wyróżnionych do rozważań stanach. Przedstawiona w artykule metodyka badania SSP ze względu na spełnienie określonych wymagań eksploatacyjnych może być użyta podczas opracowywania scenariusza pożarowego oraz projektowania systemów z uwzględnieniem różnych dostępnych rozwiązań technicznych.

1. Wprowadzenie

Transportowe systemy sygnalizacji pożaru funkcjonują w zróżnicowanych, często ekstremalnych warunkach eksploatacyjnych. Linie sygnalizacji pożaru typu A, B, pętle dozorowe, centrale alarmowe znajdują się wewnątrz obiektów budowlanych (np. dworce

kolejowe, nastawnie, stacje transformatorowe, rozdzielnie), jak i na zewnątrz, w środowisku ogólnodostępnym (np. dworce kolejowe, przejścia, wiaty magazynowe, itd.) [10,11,12]. Długookresowe badania procesu eksploatacji SSP potwierdzają tezę iż właściwe funkcjonowanie tych platform bezpieczeństwa jest funkcją niezawodności elementów składowych – czujek, modułów, central, itd. Proces obsługi i serwisu, dostęp do części zapasowych i realizacja przeglądów okresowych warunkuje także odpowiedni poziom niezawodności [4,6,16,17,19]. Analiza zjawisk eksploatacyjnych, które występują w SSP powinna uwzględniać dwa ważne zagadnienia, podejście niezawodnościowe już podczas opracowywania scenariusza projektowania zabezpieczenia pożarowego, ale także efektywność zarządzania eksploatacją tych złożonych obiektów technicznych – tj. np. dostępność serwisu, realizacji przeglądów profilaktycznych oraz parametrów środowiska w których są użytkowane [7,12,16]. W tym celu autorzy artykułu wybrali dwa reprezentatywne SSP, które są najczęściej użytkowane na rozległych terenach transportowych i dokonali odwzorowania zjawisk zachodzących w rzeczywistości eksploatacyjnej w opracowane modele badawcze [10,11].

Systemy sygnalizacji pożarowej są jednymi z bardzo istotnych elektronicznych systemów bezpieczeństwa (często instalowane w obiektach ze względu na istniejące wymagania ustawowo-prawne), które są eksploatowane na rozległych terenach transportowych. Właściwe funkcjonowanie platform bezpieczeństwa to realizacja wcześniej założonego procesu transportowego o akceptowalnym ryzyku niesprzyjających oddziaływań zewnętrznych i wewnętrznych (np. pożar, włamanie, napad, atak terrorystyczny, itd.) [6,16,19]. Zawodność poszczególnych urządzeń i systemów elektronicznych oraz błędy w działaniu operatorów nadzorujących na bieżąco proces eksploatacji mogą prowadzić do wystąpienia stanów zagrożenia lub zawodności bezpieczeństwa [5,8,18,21]. Teoria z zakresu bezpieczeństwa i ryzyka odpowiada m. in. na pytania dotyczące skutków awarii, uszkodzeń i błędów operatora. Jest to przyczyną wystąpienia stanów niedopuszczalnych w tych systemach np. zawodności lub zagrożenia bezpieczeństwa. Istotnym zagadnieniem które powinno być doprecyzowane przez eksploatatorów platform bezpieczeństwa to określenie zbioru dopuszczalnych i niedopuszczalnych stanów SSP ze względu na bezpieczeństwo danego obiektu transportowego [6,16,19,23,28,29]. Bardzo istotna jest tu kwestia prawidłowego doprecyzowania, który ze stanów SSP można uznać za dopuszczalny lub niedopuszczalny z punktu widzenia bezpieczeństwa lub opracowanego wstępnie scenariusza pożarowego dla obiektu transportowego [10,11,12,16,19].

Zbiór stanów niedopuszczalnych występujących w SSP może być odwracalny w przypadku istnienia w tym systemie elementów lub urządzeń które inicjują lub przerywają proces uszkodzenia lub awarii (w tym błędy działania operatora) [6,16,19]. Realizacja przeciwdziałania musi być wykonana w czasie dyspozycyjnym, gdzie istnieje (czas) możliwość odparowania sytuacji niebezpiecznej [16]. Takie postępowanie jest możliwe kiedy w platformach bezpieczeństwa istnieje zbiór „rezerwuar" dopuszczalnych działań przeciwawaryjnych. Wtedy nie jest możliwa realizacja przejść ze stanów dozwolonych (np. dozorowanie) do zabronionych (np. awaria modułu, centrali – stan zagrożenia bezpieczeństwa) [6,10,11,16,19].

Żywotność platform bezpieczeństwa w przypadku oddziaływań niesprzyjających można zwiększyć poprzez realizację dostępnych działań – np. stosując nadmiarowość lub rozwiązania techniczne zwiększające niezawodność samych urządzeń [10,12,16,19]. Czujka(i) które wykorzystują wiele detektorów reagujących na zjawisko pożaru. Stosowanie nadmiarowości to tolerowanie niektórych uszkodzeń a także rozbudowa systemu. Drugi przypadek to zapobieganie uszkodzeniom katastroficznym – np. czujki w systemie [10,11].

Nadmiarowość może dotyczyć zarówno samych podzespołów urządzenia, modułów systemu, jak też np. komputerów sterujących procesami eksploatacyjnymi elektronicznych

systemów bezpieczeństwa. Istotna jest też jakość informacji [6,13,14,15,16,23,24] jaką otrzymują systemy z czujników [10,11], które są zainstalowane na rozległym terenie transportowym gdzie występuje zniekształcone środowisko elektromagnetyczne (duże poziomy sygnałów zakłócających) [1,7,18,20,21]. W niektórych pracach naukowych proponuje się zastosowanie logiki rozmytej [22] lub sztucznych sieci neuronowych [2,4,5,9], które są już wykorzystywane na poziomie czujek do wypracowania sygnałów alarmu. Na funkcjonowanie transportowych systemów elektronicznych mają także istotny wpływ warunki środowiskowe, temperatura, wilgotność, drgania i wibracje [3], a także zakłócenia elektromagnetyczne [2,4,6,16,20,21] ale nie są one uwzględnione w artykule. W artykule przeprowadzono analizę eksploatacyjno-niezawodnościową SSP. Systemy te są użytkowane na rozległym terenie transportowym. Analiza otrzymanych wyników procesu eksploatacji, tj. pomiar czasów odnowy i wystąpienia uszkodzeń umożliwiła opracowanie modelu badawczego SSP, a następnie przeprowadzenie analizy niezawodnościowo-eksploatacyjnej z uwzględnieniem wyznaczonych czasów odnowy i uszkodzeń [2,8,9,16,17,19,26,28].

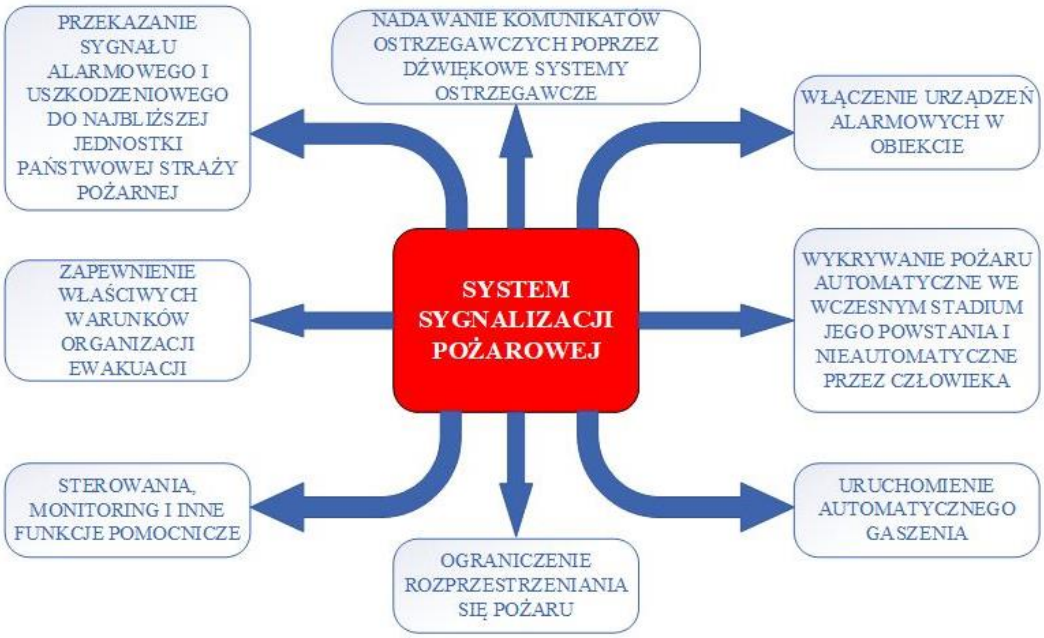
2. Reprezentatywne transportowe systemy sygnalizacji pożaru.

Obiekty transportowe w dobie szybkiego postępu technologicznego i stałego rozwoju infrastruktury narażone są na wiele zagrożeń [6,16,19,25,26,27]. Zagrożenia bezpośrednio niezwiązane z pożarem, jak np. zagrożenia terrorystyczne, mogą być jego źródłem [6,16,19]. Dlatego bardzo ważnym aspektem jest prawidłowe zabezpieczenie obiektów transportowych w czynne i bierne zabezpieczenia przeciwpożarowe – rys. 1.

Rys. 1. Statystyka liczby pożarów w obiektach obsługi pasażerów w komunikacji, w szczególności dworcach kolejowych i autobusowych, portach rzecznych i morskich, dworcach lotniczych w latach 2014-2017

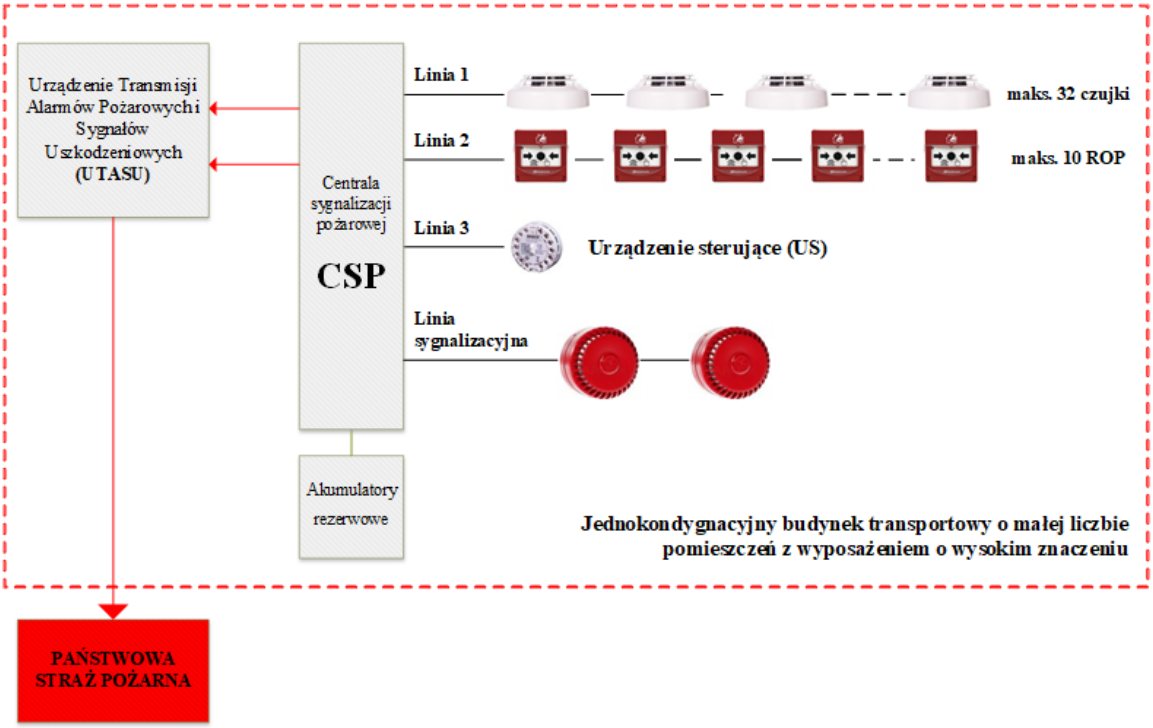
Zgodnie z rozporządzeniem Ministra Spraw Wewnętrznych i Administracji (MSWiA) Dz. U. Nr 109, poz. 719 przez techniczne środki zabezpieczenia przeciwpożarowego należy rozumieć urządzenia, sprzęt, instalacje i rozwiązania budowlane służące zapobieganiu powstawaniu i rozprzestrzenianiu się pożarów. Przez pojęcie urządzenia przeciwpożarowe podane w rozporządzeniu należy rozumieć (stałe lub półstałe, uruchamiane ręcznie lub samoczynnie urządzenia) służące do zapobiegania powstaniu, wykrywania, zwalczania pożaru lub ograniczania jego skutków. W szczególności są to stałe i półstałe urządzenia gaśnicze i zabezpieczające, urządzenia inertyzujące, urządzenia wchodzące w skład dźwiękowego systemu ostrzegawczego (DSO) i SSP. SSP to system obejmujący urządzenia

sygnalizacyjno-alarmowe, służące do samoczynnego wykrywania i przekazywania informacji o pożarze, a także urządzenia odbiorcze alarmów pożarowych i urządzenia odbiorcze sygnałów uszkodzeniowych - rysunek 2 [6,10,11,16,19].



Rys. 2. Podstawowe zadania realizowane przez system sygnalizacji pożarowej.

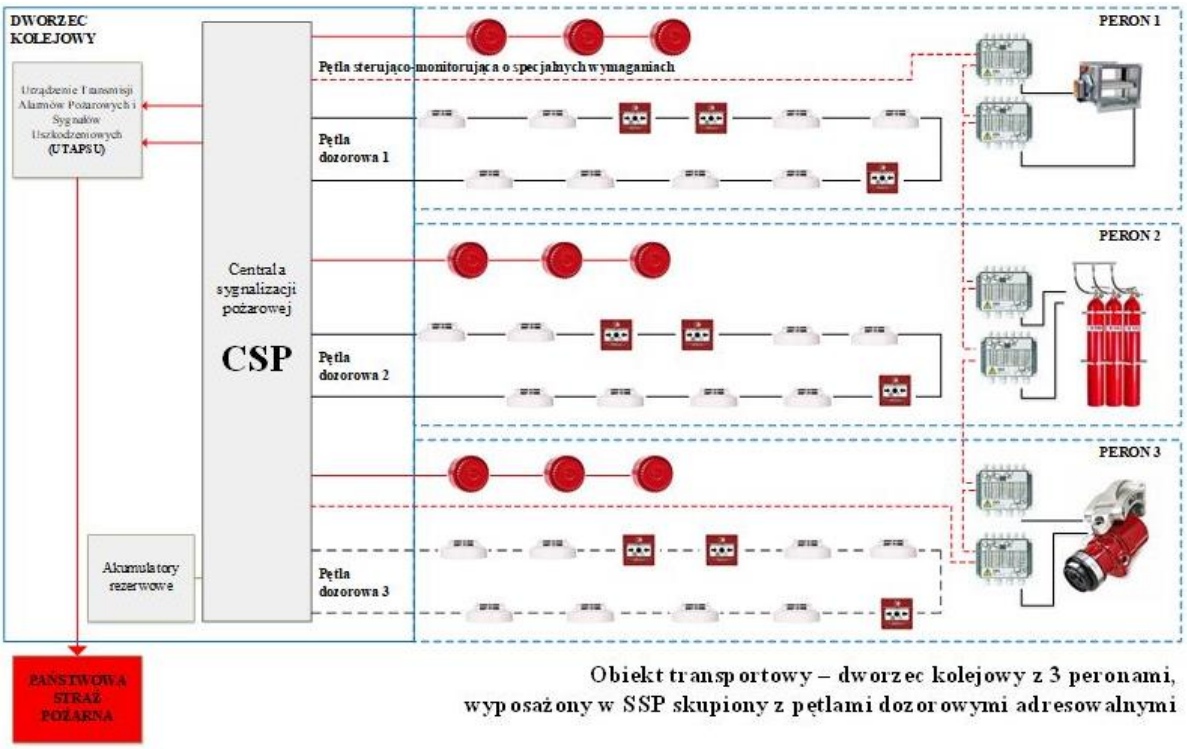
W zależności od budowy, konfiguracji oraz typu zastosowanych elementów liniowych wyróżnia się kilka rodzajów SSP - rysunek 3.



Rys. 3. SSP skupiony z liniami dozorowymi otwartymi z podłączeniem do systemu monitoringu sygnałów pożarowych i sygnałów uszkodzeniowych do PSP (CSP – centrala sygnalizacji pożaru)

Zastosowanie danego rodzaju systemu uzależnione jest od wymagań przepisów prawnych wobec SSP, scenariusza pożarowego, który musi być zrealizowany, wymagań prawnych wobec danego obiektu podlegającego ochronie, przyjętego zakresu ochrony oraz wymagań funkcjonalno-użytkowych, które ma spełniać instalacja. Dokładność wskazania miejsca powstania pożaru (źródła ognia) przez centralę zależy od zastosowanego SSP. Z kolei stawiany wymóg dokładności lokalizacji pożaru, stanowi kryterium wyboru rodzaju systemu sygnalizacji pożarowej [10,12].

W SSP konwencjonalnym (nieadresowalnym) wskazanie miejsca wykrycia pożaru jest ograniczone do linii dozorowej, natomiast w systemie adresowalnym centrala wskazuje miejsce pojawienia się pożaru z dokładnością do czujki pożarowej (w zależności od konfiguracji do strefy dozorowej) [11,12]. Rodzaj SSP, który jest zainstalowany w obiekcie transportowym ma wpływ na podział obiektu na strefy dozorowe. Pętla sterująco-monitorująca musi być wykonana wg specjalnych wymagań, w taki sposób, aby zachować ciągłość dostawy energii lub przekazu sygnału przez czas wymagany do uruchomienia i działania urządzenia, zgodnie z §187 ust. 2 Rozporządzenia Ministra Infrastruktury z dnia 12 kwietnia 2002 r. (Dz. U. Nr 75, poz. 6900 z późniejszymi zm. z uwagi na małą rozległość obiektu transportowego [10,11,12,16], krótkie odległości przebiegu okablowania pętlowego oraz małą liczbę sterowań i monitorowań, często wykorzystuje się jedną pętle sterująco-monitorująca może obsługiwać np. wszystkie perony – rysunek 4.

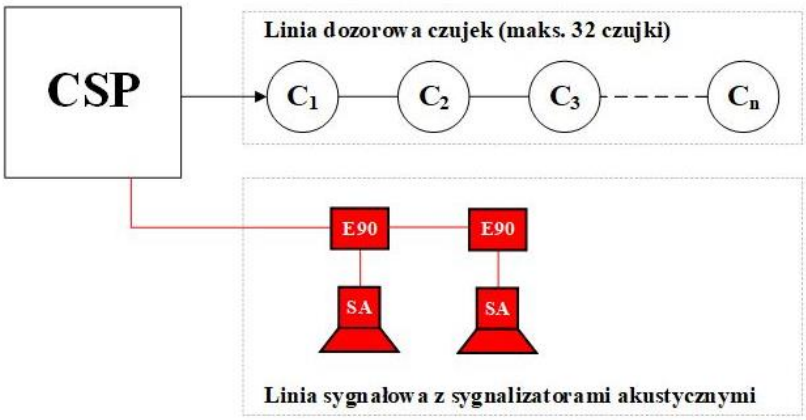


Rys. 4. Schemat systemu sygnalizacji pożarowej skupionego z pętlami dozorowymi adresowalnymi na dworcu kolejowym z trzema peronami

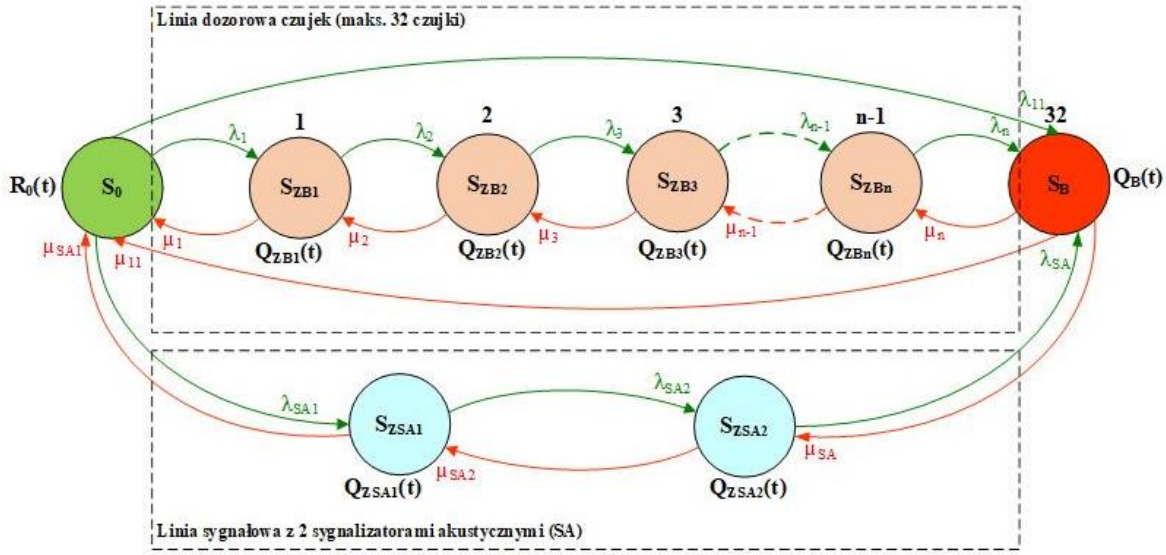
3. Analiza wybranych zagadnień procesów niezawodnościowych i eksploatacyjnych SSP

System sygnalizacji pożarowej skupiony, oparty na konwencjonalnej centrali sygnalizacji pożarowej z jedną linią dozorową otwartą wyposażoną w maksymalnie 32 czujki pożarowe oraz linią sygnałową z dwoma sygnalizatorami akustycznymi przedstawiono na rysunku 5. Natomiast na rysunku 6 przedstawiono relacje zachodzące w systemie skupionym

z centralą sygnalizacji pożarowej, do której przyłączono linię dozorową otwartą z czujkami optycznymi dymu oraz linię sygnalizacyjną z sygnalizatorami akustycznymi.



Rys. 5. SSP skupiony z linią dozorową otwartą i linią sygnałową z sygnalizatorami akustycznymi



Rys. 6. Relacje zachodzące w systemie skupionym z CSP, do której przyłączono linie dozorową otwartą z czujkami optycznymi dymu i linię sygnalizacyjną z sygnalizatorami

Relacje zachodzące w systemie – rys. 6 można opisać następującymi zależnościami (1):

$$
\begin{gathered}
R_0'(t) = -\lambda_{11} \cdot R_0(t) - \lambda_1 \cdot R_0(t) - \lambda_{SA1} \cdot R_0(t) + \mu_{11} \cdot Q_B(t) + \mu_1 \cdot Q_{ZB1}(t) + \mu_{SA1} \cdot Q_{ZSA1}(t) \\
Q'_{ZB1}(t) = -\lambda_2 \cdot Q_{ZB1}(t) - \mu_1 \cdot Q_{ZB1}(t) + \lambda_1 \cdot R_0(t) + \mu_2 \cdot Q_{ZB2}(t) \\
Q'_{ZB2}(t) = -\lambda_3 \cdot Q_{ZB2}(t) - \mu_2 \cdot Q_{ZB2}(t) + \lambda_2 \cdot Q_{ZB1}(t) + \mu_3 \cdot Q_{ZB3}(t) \\
Q'_{ZB3}(t) = -\lambda_{n-1} \cdot Q_{ZB3}(t) - \mu_3 \cdot Q_{ZB3}(t) + \lambda_3 \cdot Q_{ZB2}(t) + \mu_{n-1} \cdot Q_{ZBn}(t) \qquad (1) \\
\text{----------} \\
Q'_{ZBn}(t) = -\lambda_n \cdot Q_{ZBn}(t) - \mu_{n-1} \cdot Q_{ZBn}(t) + \lambda_{n-1} \cdot Q_{ZB3}(t) + \mu_n \cdot Q_B(t) \\
Q'_{ZSA1}(t) = -\lambda_{SA2} \cdot Q_{ZSA1}(t) - \mu_{SA1} \cdot Q_{ZSA1}(t) + \lambda_{SA1} \cdot R_0(t) + \mu_{SA2} \cdot Q_{ZSA2}(t) \\
Q'_{ZSA2}(t) = -\lambda_{SA} \cdot Q_{ZSA2}(t) - \mu_{SA2} \cdot Q_{ZSA2}(t) + \lambda_{SA2} \cdot Q_{ZSA1}(t) + \mu_{SA} \cdot Q_B(t) \\
Q'_B(t) = -\mu_{11} \cdot Q_B(t) - \mu_n \cdot Q_B(t) - \mu_{SA} \cdot Q_B(t) + \lambda_{11} \cdot R_0(t) + \lambda_n \cdot Q_{ZBn}(t) + \lambda_{SA} \cdot Q_{ZSA2}(t)
\end{gathered}
$$

Przyjmując warunki początkowe (2):

$$R_0(t) = 1$$

$$Q_{ZB1}(0) = Q_{ZB2}(0) = Q_{ZB3}(0) = \ldots = Q_{ZBn}(0) = Q_B(0) = \\ = Q_{ZSA1}(0) = Q_{ZSA2}(0) = 0 \qquad (2)$$

gdzie:

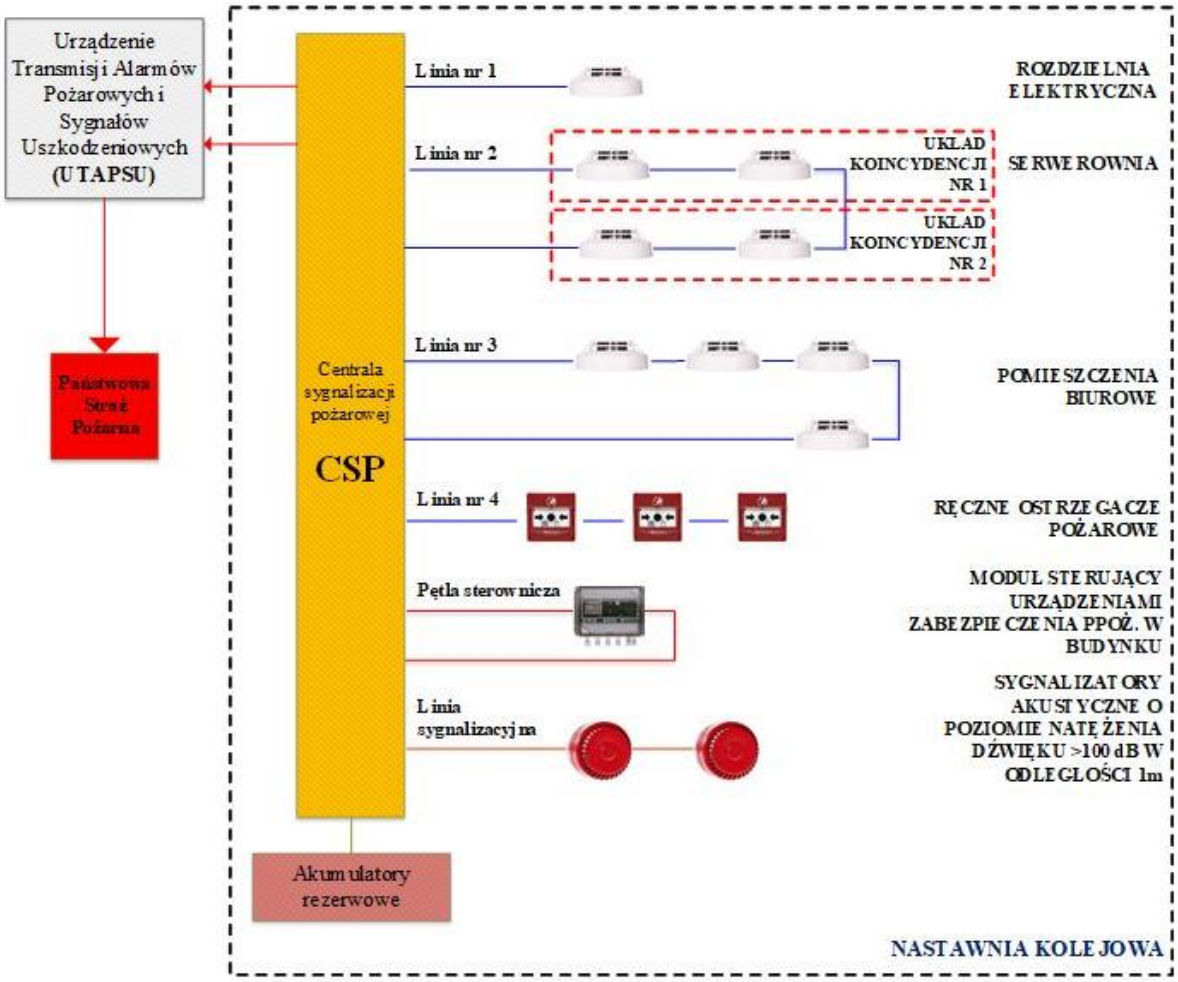
- $R_0(t)$ – funkcja prawdopodobieństwa przebywania systemu w stanie pełnej zdatności S_{PZ};
- $Q_{ZB1}(t)$, $Q_{ZBn}(t)$, $Q_{ZSA1}(t)$, $Q_{ZSA2}(t)$ – funkcja prawdopodobieństwa przebywania systemu w poszczególnych stanach zagrożenia bezpieczeństwa;
- $Q_B(t)$ – funkcja prawdopodobieństwa przebywania systemu w stanie zawodności bezpieczeństwa S_B;
- λ_{11} – intensywność przejścia ze stanu pełnej zdatności S_{PZ} do stanu zawodności bezpieczeństwa S_B;
- μ_{11} – intensywność przejścia ze stanu zawodności bezpieczeństwa S_B do stanu pełnej zdatności S_{PZ};
- λ_1, λ_1,… – intensywności przejść ze stanu pełnej zdatności S_{PZ} lub ze stanu zagrożenia bezpieczeństwa $S_{ZB1,2,\ldots}$ do stanu zawodności bezpieczeństwa $Q_B(t)$, lub stanu zagrożenia bezpieczeństwa lub zawodności bezpieczeństwa S_{ZB} – zgodnie z oznaczeniem jak na rysunku 6;
- μ_1, μ_2, … – intensywności przejść ze stanu zagrożenia bezpieczeństwa S_{ZB} do stanu pełnej zdatności S_{PZ}, ze stanu zawodności bezpieczeństwa do stanu zagrożenia bezpieczeństwa Q_{ZBn}, Q_{ZB}, Q_{ZB2}, … – zgodnie z oznaczeniami jak na rysunku 6.

Na rysunku 7 przedstawiono SSP skupiony, oparty o adresowalną CSP, do której dołączono linie otwarte czujek i ręcznych ostrzegaczy pożarowych. Wszystkie elementy wyposażone są w izolatory zwarć. System składa się z linii pętlowych, w części których zaprogramowano czujki w układach koincydencyjnych, pętli sterowniczej z modułem sterującym urządzeniami zabezpieczenia przeciwpożarowego oraz instalacjami technicznymi i bezpieczeństwa w budynku nastawni kolejowej. Do centrali podłączono także linię sygnalizacyjną z sygnalizatorami akustycznymi [6,10,11,12].

Na rysunku 7 przedstawiono relacje zachodzące w systemie w systemie skupionym z adresowalną centralą sygnalizacji pożarowej z liniami otwartymi, pętlowymi i linią sygnalizacyjną. Relacje zachodzące w systemie – rys. 7 można opisać następującymi zależnościami (3). Relacje zachodzące w systemie skupionym przedstawiono na rysunku 8.

Ze względu na różne struktury SSP, które są eksploatowane na rozległym obszarze transportowym relacje pomiędzy poszczególnymi urzadzeniami w systemach mogą być różne, co przedstawiono na rysunkach 6 i 8. System przedstawiony na rysunku 7 ma bardziej rozbudowaną strukturę niezawodnościową ze względu na występowanie większej liczby chronionych pożarowo obiektów – serwerownie, pomieszczenia biurowe i rozdzielnię elektryczną. Z tego względu można wyodrębnić osobne linie dozorowe 1,2,3 oraz pętlę sterowniczą i linie sygnalizacyjną.

Dodaktowo w linii dozorowej nr 2 zastosowano alarmowanie w układzie koincydencji. W tak zaprojektowanym SSP należy wyodrębnić więcej stanów eksploatacyjnych przez co układ równań (3) opisujących zachowanie się systemu podczas procesu eksploatacji staje się złożony.

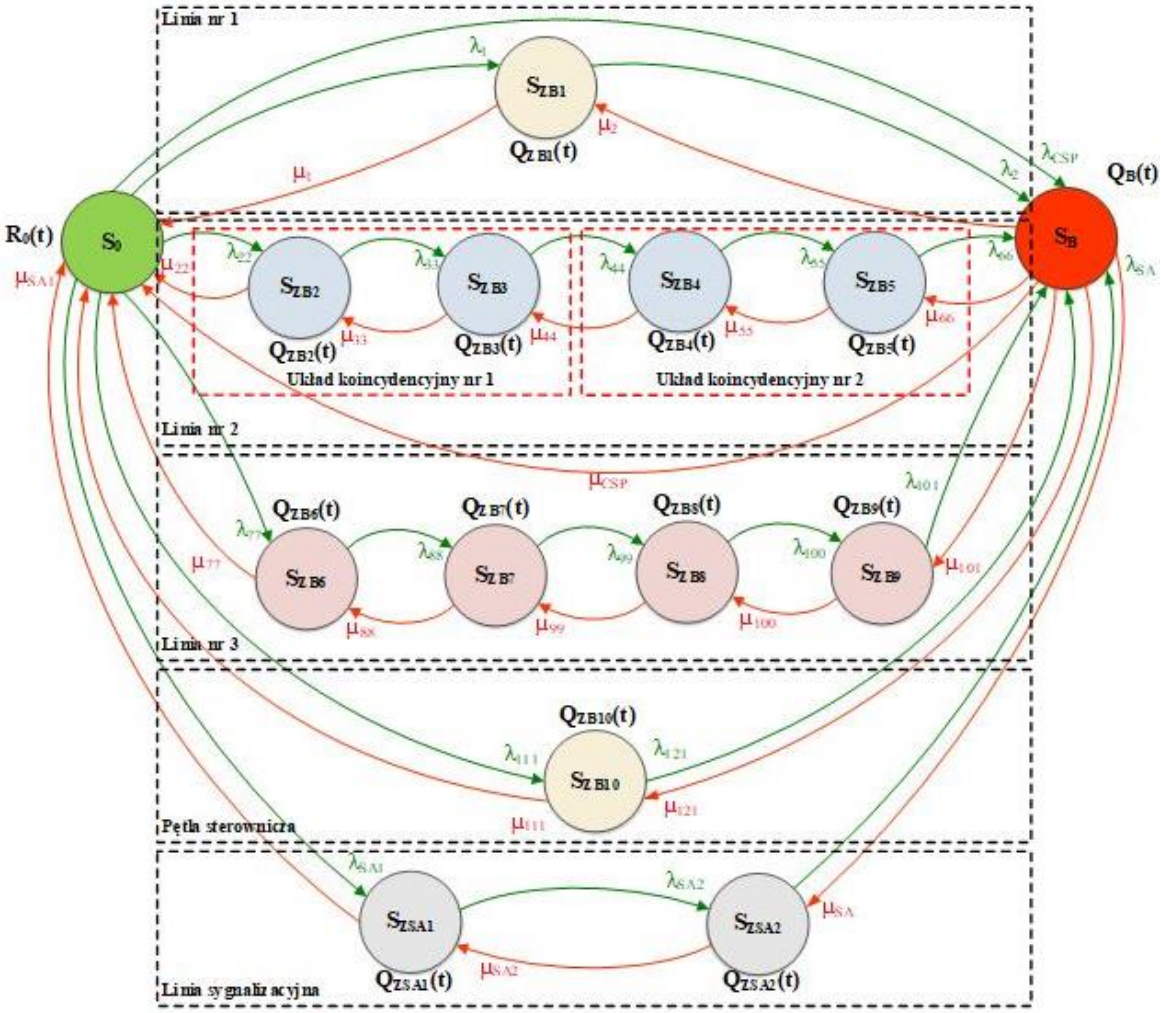


Rys. 7. SSP skupiony z adresowalną centralą sygnalizacji pożarowej z liniami otwartymi, pętlowymi i linią sygnalizacyjną

$$
\begin{aligned}
R_0'(t) &= -\lambda_{CSP}\cdot R_0(t) - \lambda_1\cdot R_0(t) - \lambda_{22}\cdot R_0(t) - \lambda_{77}\cdot R_0(t) - \lambda_{111}\cdot R_0(t) - \lambda_{SA1}\cdot R_0(t) + \\
&+ \mu_{CSP}\cdot Q_B(t) + \mu_1\cdot Q_{ZB1}(t) + \mu_{22}\cdot Q_{ZB2}(t) + \mu_{77}\cdot Q_{ZB6}(t) + \mu_{111}\cdot Q_{ZB10}(t) + \mu_{SA1}\cdot Q_{ZSA1}(t) \\
Q'_{ZB1}(t) &= -\mu_1\cdot Q_{ZB1}(t) - \lambda_2\cdot Q_{ZB1}(t) + \mu_2\cdot Q_B(t) + \lambda_1\cdot R_0(t) \\
Q'_{ZB2}(t) &= -\mu_{22}\cdot Q_{ZB2}(t) - \lambda_{33}\cdot Q_{ZB2}(t) + \mu_{33}\cdot Q_{ZB3}(t) + \lambda_{22}\cdot R_0(t) \\
Q'_{ZB3}(t) &= -\mu_{33}\cdot Q_{ZB3}(t) - \lambda_{44}\cdot Q_{ZB3}(t) + \mu_{44}\cdot Q_{ZB4}(t) + \lambda_{33}\cdot Q_{ZB2}(t) \\
Q'_{ZB4}(t) &= -\mu_{44}\cdot Q_{ZB4}(t) - \lambda_{55}\cdot Q_{ZB4}(t) + \mu_{55}\cdot Q_{ZB5}(t) + \lambda_{44}\cdot Q_{ZB3}(t) \\
Q'_{ZB5}(t) &= -\mu_{55}\cdot Q_{ZB5}(t) - \lambda_{66}\cdot Q_{ZB5}(t) + \mu_{66}\cdot Q_B(t) + \lambda_{55}\cdot Q_{ZB4}(t) \\
Q'_{ZB6}(t) &= -\mu_{77}\cdot Q_{ZB6}(t) - \lambda_{88}\cdot Q_{ZB6}(t) + \mu_{88}\cdot Q_{ZB7}(t) + \lambda_{77}\cdot R_0(t) \\
Q'_{ZB7}(t) &= -\mu_{88}\cdot Q_{ZB7}(t) - \lambda_{99}\cdot Q_{ZB7}(t) + \mu_{99}\cdot Q_{ZB8}(t) + \lambda_{88}\cdot Q_{ZB6}(t) \\
Q'_{ZB8}(t) &= -\mu_{99}\cdot Q_{ZB8}(t) - \lambda_{100}\cdot Q_{ZB8}(t) + \mu_{100}\cdot Q_{ZB9}(t) + \lambda_{99}\cdot Q_{ZB7}(t) \\
Q'_{ZB9}(t) &= -\mu_{100}\cdot Q_{ZB9}(t) - \lambda_{101}\cdot Q_{ZB9}(t) + \mu_{101}\cdot Q_B(t) + \lambda_{100}\cdot Q_{ZB8}(t) \\
Q'_{ZB10}(t) &= -\mu_{111}\cdot Q_{ZB10}(t) - \lambda_{121}\cdot Q_{ZB10}(t) + \mu_{121}\cdot Q_B(t) + \lambda_{111}\cdot R_0(t) \\
Q'_{ZSA1}(t) &= -\mu_{SA1}\cdot Q_{ZSA1}(t) - \lambda_{SA2}\cdot Q_{ZSA1}(t) + \mu_{SA2}\cdot Q_{ZSA2}(t) + \lambda_{SA1}\cdot R_0(t) \\
Q'_{ZSA2}(t) &= -\mu_{SA2}\cdot Q_{ZSA2}(t) - \lambda_{SA}\cdot Q_{ZSA2}(t) + \mu_{SA}\cdot Q_B(t) + \lambda_{SA2}\cdot Q_{ZSA1}(t) \\
Q'_B(t) &= -\mu_{CSP}\cdot Q_B(t) - \mu_2\cdot Q_B(t) - \mu_{66}\cdot Q_B(t) - \mu_{101}\cdot Q_B(t) - \mu_{121}\cdot Q_B(t) - \mu_{SA}\cdot Q_B(t) + \\
&+ \lambda_{CSP}\cdot R_0(t) + \lambda_2\cdot Q_{ZB1}(t) + \lambda_{66}\cdot Q_{ZB5}(t) + \lambda_{101}\cdot Q_{ZB9}(t) + \lambda_{121}\cdot Q_{ZB10}(t) + \lambda_{SA}\cdot Q_{ZSA2}(t)
\end{aligned}
\tag{3}
$$

Przyjmując warunki początkowe (4):

$$
\begin{gathered}
R_0(t) = 1 \\
Q_{ZB1}(0) = Q_{ZB2}(0) = Q_{ZB3}(0) = Q_{ZB4}(0) = Q_{ZB5}(0) = Q_{ZB6}(0) = Q_{ZB7}(0) = \\
= Q_{ZB8}(0) = Q_{ZB9}(0) = Q_{ZB10}(0) = Q_{ZSA1}(0) = Q_{ZSA2}(0) = Q_B(0) = 0
\end{gathered}
\tag{4}
$$



Rys. 8. Relacje zachodzące w systemie skupionym z adresowalną centralą sygnalizacji pożarowej z liniami otwartymi, pętlowymi i linią sygnalizacyjną.

4. Statystyka eksploatacyjna (naprawy, uszkodzenia) dotycząca reprezentatywnych SSP

Analizę w zakresie procesu eksploatacji SSP przeprowadzono dla n = 20 różnych systemów. Struktura badanych SSP odpowiadała reprezentatywnym systemom stosowanym do ochrony przeciwpożarowej obiektów transportowych. Badania eksploatacyjne SSP, obejmowały: odnowę, czas wystąpienie uszkodzenia i napraw. Badania zostały przeprowadzone dla następujących rodzajów SSP eksploatowanych na terenie transportowym:

a) SSP z adresowalną centralą sygnalizacji pożarowej i jedną detekcyjną pętlą dozorową (n = 15 sztuk);
b) SSP z adresowalną centralą sygnalizacji pożarowej i dwiema detekcyjnymi pętlami dozorowymi (n = 3 sztuk);

c) SSP z adresowalną CSP, z trzema pętlami dozorowymi, jedną pętlą sterująco-monitorującą do monitorowania stałych urządzeń gaśniczych oraz generowanie sygnału inicjującego do ich wyzwolenia (n = 2 sztuk).

Wszystkie wymienione SSP były eksploatowane w zbliżonych warunkach środowiskowych (temperatura, wilgotność, ciśnienie, itd.) w transportowych obiektach budowlanych. Ze względu na znaczenie SSP w zapewnieniu bezpieczeństwa w procesie transportowym, serwis zajmujący się procesem naprawy i odnowy był dostępny w ciągu 2 godzin od zgłoszenia awarii przez osoby nadzorujące eksploatację (dla n = 15 SSP). Pozostałe systemy (n = 5) miały wydłużony czas reakcji na zgłoszenie awarii do 4 godzin ze względu na nadzorowanie obiektów transportowych – budynków, które bezpośrednio nie zagrażają procesowi transportu pasażerów (np. magazyny, wiaty, itd.). W tabelach 1 - 3 przedstawiono przykładowe wyniki badań procesu eksploatacji SSP.

Tabela 1. Badanie procesu eksploatacji SSP użytkowanych w obiektach transportowych

L.p.	Rodzaj uszkodzenia	Czas wystąpienia awarii	Czas usunięcia awarii	Czas naprawy	Rodzaj naprawy
1	Zakłócenie linii nr 3	03.01.2018 godz. 14:32	03.01.2018 godz. 18:10	3h 38 min.	Poprawa przyłączenia linii do zacisków centrali i reset centrali
2	Usterka czujki 3/57	19.01.2018 godz. 8:10	19.01.2018 godz. 15:50	7h 40 min.	Reset centrali
3	Usterka czujki 3/57	01.02.2018 godz. 18:10	02.02.2018 godz. 23:30	5h 20 min.	Wymiana czujki na nową
.....					
n-2	Zakłócenie linii nr 1	01.12.2018 godz. 4:15	01.12.2018 godz. 9:00	4h 45 min.	Reset centrali
n-1	Usterka czujki 3/11	15.12.2018 godz. 11:15	15.12.2018 godz. 14:20	4h 5 min.	Wymiana gniazda czujki
n	Błąd Cos2 komunikacji	27.12.2018 godz. 15:05	28.12.2018 godz. 9:05	18h 5 min.	Reset CSO2 i centrali sygnalizacji pożarowej

W tabeli 1 przedstawiono reprezentatywne rodzaje uszkodzenia dla wybranych SSP. Dane opracowano na podstawie zbioru uszkodzeń z n = 20 SSP, które są eksploatowane na rozległym obszarze transportowym. Dla danego rodzaju uszkodzenia występującego w SSP (n = 20 sztuk), przyjęto maksymalny czas naprawy. W czasie naprawy nie uwzględniono czasów przyjazdu serwisu (dla tego rodzaju SSP serwis powinien być na miejscu).

Tabela 2. Rodzaje napraw wraz z oznaczeniem maksymalnego czasu T_{max} w ujęciu rocznym

L.p.	Naprawa z danego rodzaju uszkodzenia	Czas wystąpienia awarii	Czas usunięcia awarii	Maksymalny czas naprawy [T_{max}]
Uszkodzenie pętli dozorowej 1				
1.	Zakłócenie linii nr 3	03.01.2018 godz. 14:32	03.01.2018 godz. 18:10	3h 38 min.
2.	Zakłócenie linii nr 2	11.03.2018 godz. 15:00	11.03.2018 godz. 16:30	1h 30 min.
3.	Doziemienie pętli nr 1	02.05.2018 godz. 13:30	02.05.2018 godz. 19:00	5h 30 min.
4.	Zakłócenie linii nr 1	01.12.2018 godz. 4:15	01.12.2018 godz. 9:00	4h 45 min.
5.	Błąd komunikacji pętla 1	30.11.2018 godz. 10:30	30.11.2018 godz. 14:30	4h
Uszkodzenie ręcznego ostrzegacza pożarowego				
1.	Usterka ROP 1/10	15.06.2018 godz. 9:20	15.06.2018 godz. 13:20	4h
2.	Usterka ROP 1/10	16.06.2018 godz. 14:00	16.06.2018 godz. 19:05	5h 5 min.
……**Uszkodzenie zasilania SSP**……..				
1.	Usterka zasilania 230V	27.02.2018 godz. 11:30	27.02.2018 godz. 11:45	15 min.
2.	Usterka akumulatorów CSP	16.04.2018 godz. 19:00	17.04.2018 godz. 8:10	13h 10 min.

Tabela 3. Intensywność uszkodzeń wraz z oznaczeniem czasów niezdatności systemu sygnalizacji pożarowej w ujęciu rocznym (przykład).

L.p.	Naprawa z danego rodzaju uszkodzenia	Czas wystąpienia awarii	Czas usunięcia awarii	Czas naprawy [T_{max}]
Uszkodzenie pętli dozorowej				
1.	Zakłócenie linii nr 3	03.01.2018 godz. 14:32	03.01.2018 godz. 18:10	3h 38 min.
2.	Doziemienie pętli nr 1	02.05.2018 godz. 13:30	02.05.2018 godz. 19:00	5h 30 min.
3.	Zakłócenie linii nr 1	01.12.2018 godz. 4:15	01.12.2018 godz. 9:00	4h 45 min.
4.	Błąd komunikacji pętla 1	30.11.2018 godz. 10:30	30.11.2018 godz. 14:30	4h
Sumaryczny czas niezdatności SSP w ujęciu rocznym:				19h 23 min.
Uszkodzenie zasilania SSP				
1.	Usterka zasilania 230V	27.02.2018 godz. 11:30	27.02.2018 godz. 11:45	15 min.
2.	Usterka akumulatorów CSP	16.04.2018 godz. 19:00	17.04.2018 godz. 8:10	13h 10 min.
Sumaryczny czas niezdatności SSP w ujęciu rocznym:				13h 25 min.

5. Modelowanie procesu eksploatacji SSP w programie RELIASOFT BLOCKSIM

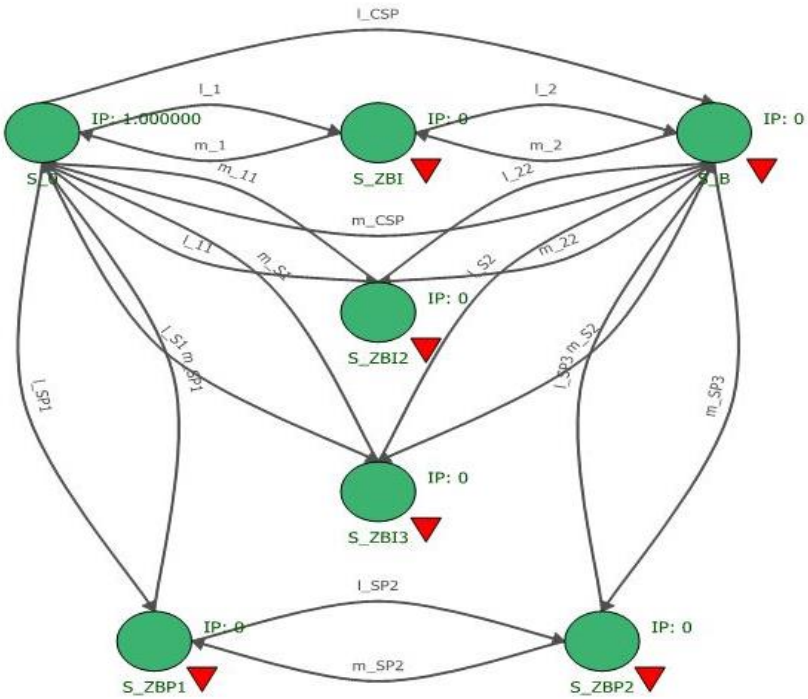
Obliczenia prawdopodobieństwa przebywania systemu w stanach zagrożenia bezpieczeństwa, zawodności bezpieczeństwa, stanie pełnej zdatności dla modelu procesu eksploatacji SSP przeprowadzono w komercyjnym, specjalistycznym programie obliczeniowym firmy ReliaSoft typu BlockSim. Obliczenia przeprowadzono dla modelu SSP skupionego – linie otwarte, bez powiadamiania. W tabelach 4 i 5 przedstawiono obliczone parametry – np. prawdopodobieństwo początkowe i średnie, współczynnik gotowości dla czasu t w poszczególnych stanach, czas w jakim SSP znajduje się w danym stanie.

Tabela 4. Parametry systemu sygnalizacji pożarowej dla czasu t = 8760 [h]

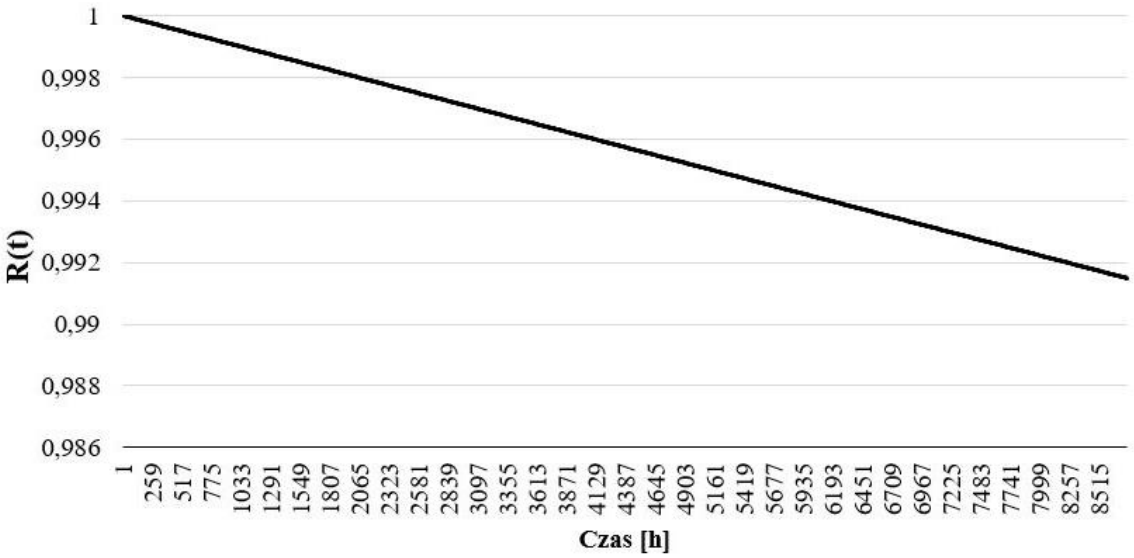
Nazwa stanu	Prawdopodobieństwo początkowe	Średnie prawdopodobieństwo	Gotowość dla czasu t [8760 h]	Nieuszkadzalność dla czasu t	Czas spędzony w danym stanie
S_0	1	0,999993444	0,999993439	0,991489928	8759,94257
S_B	0	2,245 E-07	2,24528 E-07	0,001526641	0,001966621
S_{ZBI}	0	3,75408 E-06	3,75731 E-06	0,003920964	0,032885763
S_{ZBI2}	0	8,25865 E-07	8,26355 E-07	0,001033117	0,00723458
S_{ZBI3}	0	7,10979 E-07	7,11386 E-07	0,000996234	0,006228174
S_{ZBP1}	0	8,16121 E-07	8,16726 E-07	0,001033117	0,007149221
S_{ZBP2}	0	2,24374 E-07	2,24529 E-07	0	0,001965516

Tabela 5. Macierz intensywności przejść dla poszczególnych stanów SSP dla t = 8760 [h]

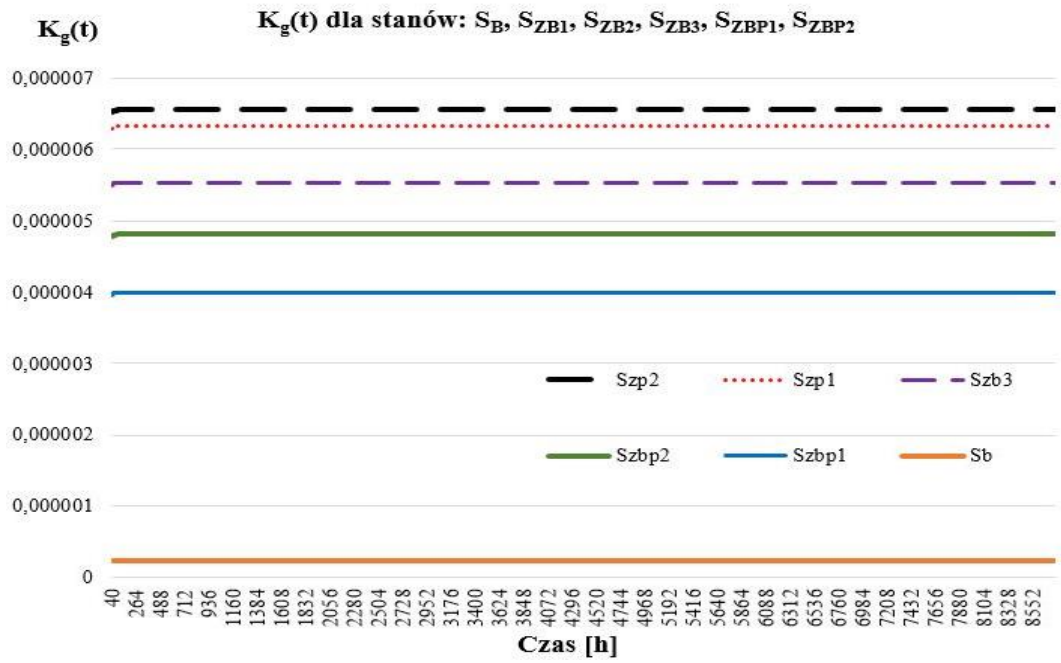
Z → do	**S_0**	**S_B**	**S_{ZBI}**	**S_{ZBI2}**	**S_{ZBI3}**	**S_{ZBP1}**	**S_{ZBP2}**
S_0	-	1,7502 E-07	4,49514 E-07	1,1844 E-07	1,14212 E-07	1,1844 E-07	0
S_B	0,0759	-	0,1818	0,1968	0,125	0	0,2
S_{ZBI}	0,1305	2,52906 E-07	-	0	0	0	0
S_{ZBI2}	0,1968	5,70919 E-08	0	-	0	0	0
S_{ZBI3}	0,2	1,4161 E-08	0	0	-	0	0
S_{ZBP1}	0,2	0	0	0	0	-	1,18 E-07
S_{ZBP2}	0	1,4161 E-08	0	0	0	0,2	-



Rys. 9. Migracja możliwych stanów SSP skupionego z liniami otwartymi, bez powiadamiania PSP (gdzie: PSP – Państwowa Straż Pożarna).



Rys. 10. Nieuszkadzalność R(t) systemu SSP z liniami otwartymi, bez powiadamiania PSP

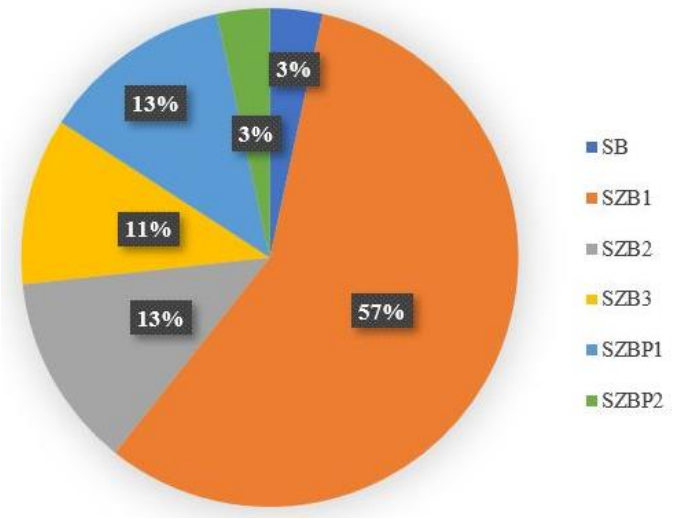


Rys. 11. Współczynniki gotowości strefowe (cząstkowe) dla stanów S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2} SSP (zawodność i zagrożenie bezpieczeństwa); na wykresie nie zobrazowano stanu S_0 (dla t = 0 $S_0(t) = 1$)

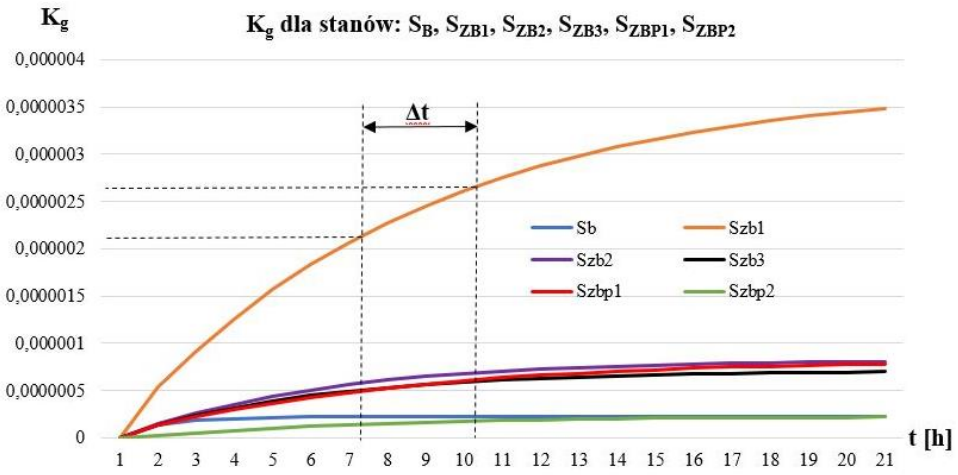
Dla przykładowego czasu eksploatacji t = 4 201 h wartości współczynnika gotowości $K_g(t)$ dla poszczególnych stanów S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2} systemu sygnalizacji pożarowej przedstawiono w tabeli 6, a udział procentowy przebywania SSP w określonym stanie na rysunku 12. Na rysunku 13 przedstawiono szybkość narastania wskaźników gotowości stanów dla wybranego przedziału czasu eksploatacji SSP.

Tab. 6. Wartości współczynnika $K_g(t)$ dla poszczególnych stanów SSP w czasie

Czas [h]	Stany systemu sygnalizacji pożarowej					
	S_B	S_{ZB1}	S_{ZB2}	S_{ZB3}	S_{ZBP1}	S_{ZBP2}
	Wartość współczynnika $K_g(t)$					
4201	$2245282 \bullet 10^{-6}$	$3757312 \bullet 10^{-5}$	$82635535 \bullet 10^{-6}$	$71138635 \bullet 10^{-6}$	$81672633 \bullet 10^{-6}$	$2245287 \bullet 10^{-6}$



Rys. 12. Procentowy udział przebywania SSP w określonym stanie w zgodnie z tabelą 6



Rys. 13. Szybkość narastania wskaźników gotowości stanów S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2} systemu sygnalizacji pożarowej (zawodność i zagrożenie bezpieczeństwa); na wykresie nie zobrazowano stanu S_0 (dla t = 0, $K_g(t) = 1$), przyjęto czas t = 21 h w celu zobrazowania prędkości zmian parametrów w początkowej fazie zmian stanów przejść

Szybkość narastania wskaźnika gotowości strefowej dla określonego stanu można opisać za pomocą wzoru (5):

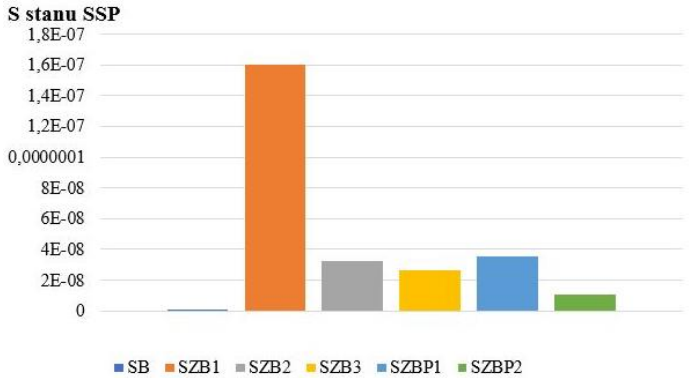
$$S_{ZB1} = \frac{\Delta Kg_{SZB1}}{\Delta t}\left[\frac{1}{h}\right] \tag{5}$$

$$S_{ZB1} = \frac{(2{,}75466E\text{-}6) - (2{,}27438E\text{-}6)}{10{,}25 - 7{,}25} = 1{,}08915E\text{-}10\left[\frac{1}{h}\right]$$

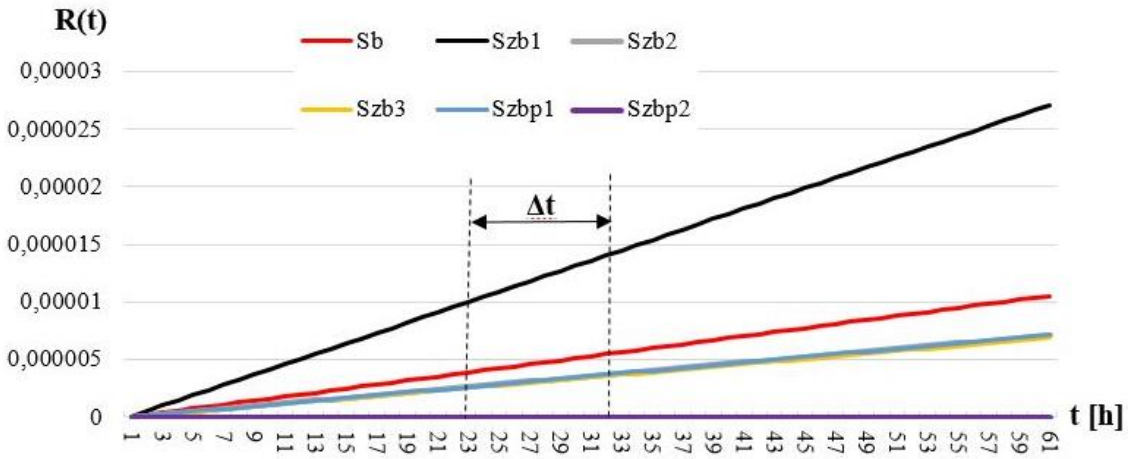
Wartości szybkości narastania wskaźnika gotowości K_g dla pozostałych stanów systemu sygnalizacji pożarowej przedstawiono w tabeli 7 oraz na rysunku 14. Na rysunku 15 przedstawiono prawdopodobieństwo przebywania SSP w poszczególnych stanach.

Tab. 7. Szybkość narastania wskaźników gotowości K_g dla poszczególnych stanów SSP.

L.p.	Stan SSP	Szybkość S narastania wskaźnika gotowości K_g [1/h]
1.	S_B	1,37487E-10
2.	S_{ZB1}	2,08915E-10
3.	S_{ZB2}	4,99493E-11
4.	S_{ZB3}	2,2907E-11
5.	S_{ZBP1}	1,05053E-11
6.	S_{ZBP2}	4,81767E-12



Rys. 14. Szybkość narastania wskaźników gotowości stanów S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2}

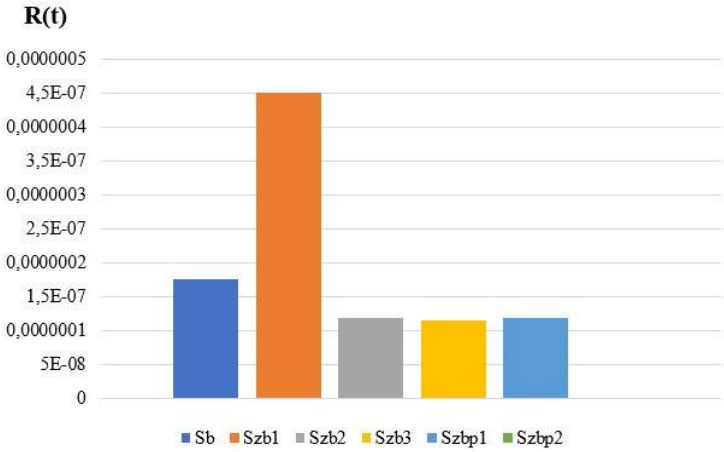


Rys. 15. Prawdopodobieństwo przebywania SSP w stanie R(t) dla stanów S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2}; na wykresie nie zobrazowano R(t) dla stanu S_0 (dla t = 0, R(t) = 1), przyjęto czas t = 61 h w celu zobrazowania prędkości zmian parametrów w początkowej fazie zmian

Szybkość narastania R(t) w czasie Δt dla określonego stanu opisujemy za pomocą wzoru (6):

$$S_{ZB1} = \frac{\Delta R(t)_{SZB1}}{\Delta t}\left[\frac{1}{h}\right] \tag{6}$$

$$S_{ZB1} = \frac{(2{,}75466E-6)-(2{,}27438E-6)}{10{,}25-7{,}25} = 1{,}08915E-10\left[\frac{1}{h}\right]$$



Rys. 16. Szybkość narastania wartości R(t) w czasie dla wybranych stanów SSP

6. Wnioski

Systemy sygnalizacji pożaru funkcjonujące na rozległych terenach transportowych posiadają różne struktury połączeń, które są funkcją wykonywanych zadań – dozorowanie pożarowe obiektów budowlanych [6,10,11,12,16]. Złożone SSP posiadają kilka – kilkanaście linii pętlowych dozorujących, linie: sygnalizatorów, sterowania oddymianiem, gaszenia gazem, itd. Ze względu na zakres wykonywanych zadań i sterowań pożarowych struktura niezawodnościowo-eksploatacyjna takich systemów jest mieszana. Stosuje się wszystkie dostępne środki technicznie celem zwiększenie niezawodności SSP. W artykule przedstawiono model i analizę eksploatacyjno-niezawodnościową wybranego SSP, który jest użytkowany na terenie transportowym. W systemie wyróżniono siedem stanów

eksploatacyjnych. Średnia wartość prawdopodobieństwa przebywania systemu w stanie zdatności wynosi S_0 = 0,999993444, natomiast czas spędzony w tym stanie to 8759,94257 [h] (symulację przeprowadzono dla t = 1 rok użytkowania SSP). Rozpatrując tzw. współczynniki gotowości strefowe (cząstkowe) $K_{gs}(t)$ dla stanów S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2} SSP można zauważyć iż w początkowym czasie eksploatacji dominuje stan S_{ZBP2}.

Dlatego projektując SSP należy zwrócić szczególną uwagę na przejście pomiędzy stanami zdatności S_0 a stanem zagrożenia bezpieczeństwa S_{ZBP2}. Określono szybkość narastania wskaźników gotowości stanów S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2} SSP (zawodność i zagrożenie bezpieczeństwa) w celu zobrazowania szybkości zmian parametrów w początkowej fazie zmian stanów przejść. W początkowym czasie eksploatacji SSP największą wartość otrzymano dla stanu S_{ZB1} = 2,08915E-10 [1/h]. Wszystkie współczynniki gotowości strefowej (tzw. cząstkowe) w dalszym procesie eksploatacji stabilizują swoje wartości na stałych poziomach – rys. 13. Wartość prawdopodobieństwa przebywania SSP w stanie R(t) jest bardzo mała dla poszczególnych stanów S_B, S_{ZB1}, S_{ZB2}, S_{ZB3}, S_{ZBP1}, S_{ZBP2} w początkowym okresie eksploatacji - rys. 15. Największa szybkość narastania wartości R(t) w czasie początkowego procesu eksploatacji była dla stanu S_{ZB1} SSP.

Bibliografia

1. Białek K, Paś J. Exploitation of selected railway equipment - conducted disturbance emission examination, Diagnostyka 2018; 19(3): 29-35.

2. Branson D. Stirling numbers and Bell numbers, their role in combinatorics and probability, Math. Scientist 2000; (25): 1-31.

3. Burdzik R, Konieczny Ł, Figlus T. Concept of on-board comfort vibration monitoring system for vehicles, In the monograph Activities of Transport Telematics Springer 2013; 418-425.

4. Duer S, Zajkowski K, Duer R, Paś J. Designing of an effective structure of system for the maintenance of a technical object with the using information from an artificial neural network, Neural Computing & Applications 2012; 23(3): 913–925.

5. Duer S, Scaticailov S, Paś J, Duer R, Bernatowicz D. Taking decisions in the diagnostic intelligent systems on the basis information from an artificial neural network, 22nd International Conference on Innovative Manufacturing Engineering and Energy MATEC Web of Conferences 2018; (178): 1-6.

6. Dyduch J, Paś J, Rosiński A. The basic of the exploitation of transport electronic systems, Radom Publishing House of Radom University of Technology 2011.

7. Dziula P, Paś J. Low Frequency Electromagnetic Interferences Impact on Transport Security Systems Used in Wide Transport Areas, TransNav the International Journal on Marine Navigation and Safety of Sea Transportation 2018; 12(2): 251-258.

8. Garmabaki AHS, Ahmadi A, Mahmood YA, Barabadi A. Reliability modelling of multiple repairable units, Quality and Reliability Engineering International 2016; 32(7): 2329–2343.

9. Jachimowski R, Żak J, Pyza D. Routes planning problem with heterogeneous suppliers demand, 21st International Conference on Systems Engineering Las Vegas USA 2011; 434-437.

10. Klimczak T, Paś J. Analysis of reliability structures for fire signaling systems in the field of fire safety and hardware requirements, Journal of KONBIN 2018; (64): 191-214.

11. Klimczak T, Paś J. Electromagnetic environment on extensive logistic areas and the proces of using electronic safety system, Politechnika Warszawska Prace Naukowe Transport 2018; (121): 135-146.

12. Klimczak T, Paś J. Analysis of solution of a fire signaling system for a choice railway building, Biuletyn WAT 2018; (67)4: 195-205.

13. Krzykowski M, Paś J, Rosiński A. Assessment of the level of reliability of power supplies of the objects of critical infrastructure, IOP Conf. Series Earth and Environmental Science 2019; 1-9.

14. Lewiński A, Perzyński T, Toruń A. The analysis of open transmission standards in railway control and management, Communications in Computer and Information Science, Berlin Heidelberg Springer-Verlag 2012; (329): 10-17.

15. Łubkowski P, Laskowski D. Selected issues of reliable identification of object in transport systems using video monitoring services, Communication in Computer and Information Science, Berlin Heidelberg Springer 2015; (471): 59-68.

16. Paś J. Operation of electronic transportation systems, Radom Publishing House of University of Technology and Humanities, 2015.

17. Paś J, Duer S. Determination of the impact indicators of electromagnetic interferences on computer information systems, Neural Computing & Applications 2012; 23(7): 2143-2157.

18. Paś J, Rosiński A. Selected issues regarding the reliability-operational assessment of electronic transport systems with regard to electromagnetic interference, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19(3): 375–381.

19. Rosiński A. Modelling the maintenance process of transport telematics systems, Warsaw Publishing House of Warsaw University of Technology, 2015.

20. Siergiejczyk M, Paś J, Rosiński A. Train call recorder and electromagnetic interference, Diagnostyka 2015; 16(1): 19-22.

21. Siergiejczyk M, Paś J, Rosiński A. Issue of reliability–exploitation evaluation of electronic transport systems used in the railway environment with consideration of electromagnetic interference, IET Intelligent Transport Systems 2016; 10(9): 587–593.

22. Skorupski J, Uchroński P. A fuzzy reasoning system for evaluating the efficiency of cabin luggage screening at airports, Transportation Research Part C - Emerging Technologies 2015; (54): 157-175.

23. Stawowy M. Model for information quality determination of teleinformation systems of transport, In: Proceedings of the European Safety and Reliability Conference ESREL 2014 CRC Press/Balkema 2015; 1909–1914.

24. Stawowy M, Kasprzyk Z. Identifying and simulation of status of an ICT system using rough sets, Tenth International Conference on Dependability and Complex Systems DepCoS-RELCOMEX Springer 2015; (365): 477-484.

25. Warczek J, Młyńczak J, Celiński I. Simulation studies of a shock absorber model proposed under conditions of different kinematic input functions, Vibroengineering Procedia 2015; 6: 248–253.

26. Weintrit A, Dziula P, Siergiejczyk M, Rosiński A. Reliability and exploitation analysis of navigational system consisting of ECDIS and ECDIS back-up systems, The monograph Activities in Navigation - Marine Navigation And Safety Of Sea Transportation London CRC Press/Balkema 2015; 109-115.

27. Weintrit A. Technical infrastructure to support seamless information exchange in e-Navigation, TST Springer Heidelberg CCIS 2013; (395): 188–199.

28. Yang L, Yan X. Design for Reliability of Solid State Lighting Products, Solid State Lighting Technology and Application Series Springer New York 2013; (1): 497-556.

29. Zajkowski K, Rusica I, Palkova Z. The use of CPC theory for energy description of two nonlinear receivers, MATEC Web of Conferences 2018; (178): 1-6.

dr hab. inż. Barbara Kucharska

Department of Production Engineering and Materials Technology
Czestochowa University of Technology
Armii Krajowej 19 str., 42-200 Czestochowa, Poland
E-mail: kucharska.barbara@wip.pcz.pl

Identyfikacja naprężeń powierzchniowych w rurze do układu wydechowego wykonanej technologią hydroformowania na podstawie pomiarów dyfraktometrycznych

Słowa kluczowe: *hydroformowanie, układ wydechowy, naprężenia powierzchniowe, rentgenowski pomiar naprężeń*

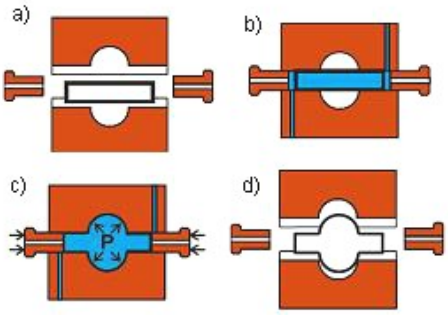
Streszczenie: W pracy dokonano identyfikacji naprężeń powierzchniowych w rurze wydechowej ze stali Cr-Ni kształtowanej technologią hydroformowania. Naprężenia wyznaczono nieniszczącą rentgenowską metodą $\sin^2\psi$. Na powierzchni rury stwierdzono złożony stan naprężeń rozciągających o wartościach z zakresu 69-240 MPa dla naprężeń obwodowych i 26-290 MPa dla naprężeń wzdłużnych. Rozłożenie naprężeń na obwodzie i długości rury analizowano na podstawie współczynników zmienności i grubości ścianki. Stwierdzono zależność pomiędzy wartością naprężeń powierzchniowych a grubością ścianki rury. Największe naprężenia występowały w obszarach rury gdzie grubość ścianki była najsilniej zredukowana. W centralnej części rury gdzie redukcja grubości ścianki była najmniejsza naprężenia również były najmniejsze, ale cechowały się największym rozproszeniem wartości.

1. Wprowadzenie

Do produkcji układów wydechowych w przemyśle samochodowym obecnie najczęściej stosowane są blachy ze stali ferrytycznej pokryte powłokami ochronnymi ze stopów aluminium. Aluminiowe powłoki zapewniają ochronę stali przed oddziaływaniem korozyjnego medium w podwyższonej temperaturze, w tym spalin, a także odporność na ścieranie w przypadku ogniowych powłok AlSi [18, 19, 13]. Obecnie, w produkcji układów wydechowych coraz bardziej preferowane jest stosowanie stali austenitycznych [3,4]. Stale te wykazują wyjątkowo korzystną kombinację właściwości chemicznych i możliwości plastycznego kształtowania [4, 6, 17]. Światowa produkcja austenitycznych stali utrzymuje się na wysokim poziomie z ciągłą tendencją wzrostową. Około 95% produkcji austenitycznych stali odpornych na korozję stanowią produkty kształtowane plastycznie, z czego prawie 10% stosowane jest w motoryzacji [6, 8]. Dobra odkształcalność plastyczna stali austenitycznych rekompensuje zarówno ich wyższą cenę, jak i stosowanie kosztownych technologii, takiej jak na przykład hydroformowanie.

Hydroformowanie jest metodą kształtowania arkuszy blach lub profili zamkniętych z użyciem płynu (najczęściej wody) pod ciśnieniem (Rys.1) [11, 1]. Zaletą metody są

zmniejszenie liczby spawanych połączeń w konstrukcjach i uzyskanie części o lepszym stanie powierzchni, cieńszych ścianach i lepszej tolerancji wymiarowej [14, 15, 5]. Obecnie, kształtowanie profili zamkniętych metodą hydroformowania jest szczególnie popularne w produkcji rowerów (ramy aluminiowe) i w przemyśle motoryzacyjnym. Metodą hydroformowania produkuje się już nadwozia samochodów, ramy nośne, tłumiki i inne części, w tym detale konstrukcyjne układów wydechowych [9].



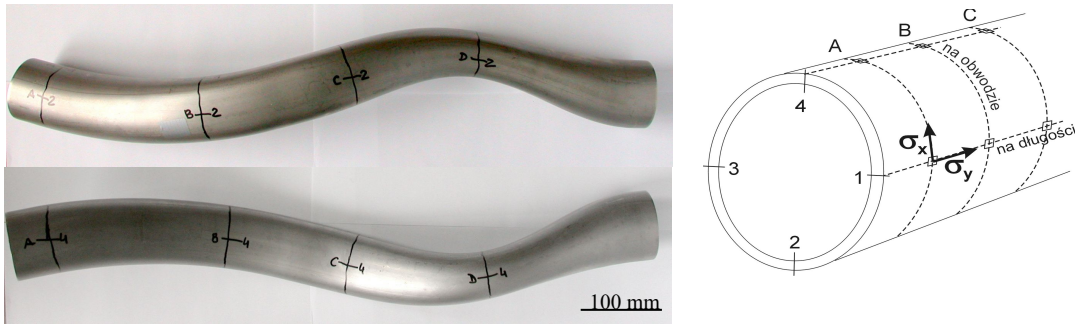
Rys. 1. Schemat operacji hydroformowania rury: a) umieszczenie rury w matrycy, b) mocowanie i uszczelnianie końców rur i wprowadzanie płynu do wnętrza rury, c) zamknięcie wlotu płynu i kształtowanie rury pod ciśnieniem płynu, d) wypuszczenie płynu i pobieranie uformowanej części z matrycy

Metodą hydroformowania możliwe jest uzyskanie złożonych kształtów detali, ze zróżnicowaną krzywizną, trudnych do uzyskania tradycyjnymi metodami przeróbki plastycznej [3, 9, 10]. Jest to niezwykle istotne z punktu widzenia konieczności upakowania wielu części mechaniki samochodowej w jak najmniejszej przestrzeni. Jednocześnie, w przypadku profili zamkniętych muszą one zapewniać swobodny przepływ mediów, np. spalin.

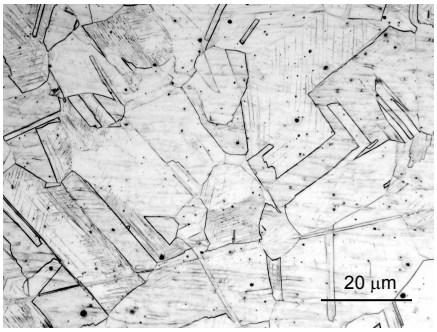
Szczególne warunki hydromechanicznego kształtowania rur, w których materiał nie ma możliwości swobodnego „płynięcia" w obszarze krawędzi, jak w przypadku blach, powodują powstawanie w materiale dużych naprężeń [7, 20]. Wysoki poziom naprężeń w detalu urządzenia sprawia, że jest on podatny na niestabilność wymiarową. Co więcej, nawet niewielkie uszkodzenia mechaniczne lub korozyjne, zainicjowane podczas eksploatacji takiego detalu, będą powodowały nieproporcjonalnie duże jego odkształcenia w wyniku relaksacji naprężeń. Z punktu widzenia trwałości eksploatacyjnej istotnym jest zatem określenie tych naprężeń metodami nieniszczącymi i ich zniwelowanie. W niniejszej pracy przedstawiono wyniki pomiaru naprężeń w rurze do układu wydechowego wykonanej technologią gięcia i hydroformowania.

2. Materiał i metodyka badań

Badano rurę przeznaczoną do samochodowego układu wydechowego, której w procesie produkcji finalny kształt nadano technologią hydroformowania (Rys. 2). Rura poddana hydroformowaniu, o grubości ścianki max. 1,7 mm, była wykonana ze stali chromowo-niklowej gatunku X2CrNi18-9 (AISI 304L) o mikrostrukturze austenitycznej (Rys. 3). Stężenie niklu w stali określone spektroskopowo wynosiło 9.6%wag., co zapewnia stali lepszą ciągliwość (DDQ – deep draw quality) w porównaniu do standardowej wersji gatunku 18-8.



Rys. 2. a) Widok ogólny rury poddanej badaniom i b) schemat oznaczenia miejsc pomiarów naprężeń (□) na powierzchni rury. Oznaczenia A, B, C i D – obwód rury, 1, 2, 3 i 4 – punkty na obwodzie rury usytuowane co 90°



Rys. 3. Mikrostruktura stali na przekroju poprzecznym rury

Celem badań było określenie naprężeń na powierzchni zewnętrznej rury w kierunku obwodowym (x) i wzdłużnym (y). Na potrzeby badań na powierzchni rury wytypowano cztery obszary (obwody), spośród których trzy (oznaczone jako B, C i D – Fig. 2a) znajdowały się w miejscach, w których kształt rury uległ największej zmianie i jeden (oznaczony jako A – Fig. 2a), blisko końca rury, gdzie przekrój poprzeczny rury był najbardziej zbliżony do kołowego. Naprężenia wyznaczono w czterech punktach na każdym z wybranych obwodów rury (oznaczone jako 1, 2, 3 i 4 – Fig. 2b), rozmieszczonych co ~90° w ten sposób, że np. punkty A1, B1, C1 i D1, układały się wzdłuż rury na jednej tworzącej.

Do wyznaczenia naprężeń zastosowano metodę dyfrakcji rentgenowskiej, znaną jako metodę $\sin^2\psi$ [16, 2]. Badania wykonano w oparciu o preferowany do pomiarów naprężeń w stalach austenitycznych refleks dyfrakcyjny od płaszczyzny (311) [12]. Pomiary wykonano z użyciem dyfraktometru PROTO dedykowanego do pomiarów naprężeń w Katedrze Nauki o Materiałach na Wydziale Budowy Maszyn i Lotnictwa w Politechnice Rzeszowskiej (Fig. 4).

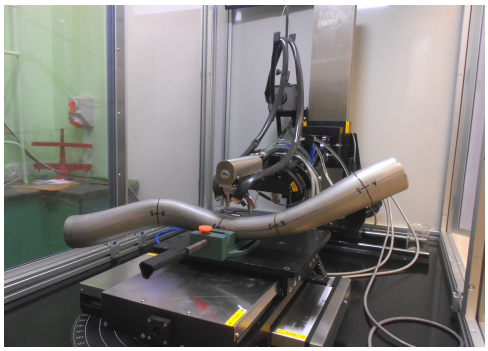


Fig. 4. Pomiar naprężeń w hydroformowanej rurze z użyciem dyfraktometru PROTO

Zastosowano promieniowanie $K_{\alpha}Mn$ (kolimator ∅2mm) o długości 0,2103 nm, które umożliwiło pomiar naprężeń w warstwie przypowierzchniowej stali o maksymalnej grubości ok. 17 μm.

Wyznaczenie naprężeń metodą rentgenowską polega na określeniu odkształcenia sieci krystalicznej ε spowodowanego m.in. przeróbką plastyczną materiału polikrystalicznego. Odkształcenie to zdefiniowane jest jako względna różnica odległości międzypłaszczyznowych Δd w materiale z naprężeniami i bez naprężeń. Naprężenia σ_{ϕ} wylicza się z zależności (1), w której ϕ oznacza kierunek naprężeń (określony poprzez usytuowanie detalu w trakcie pomiaru), natomiast ψ - kąt ustawienia głowicy dyfraktometru lub nachylenia powierzchni detalu przy pomiarze d_{hkl} sieci odkształconej.

$$\varepsilon_{\phi\psi} = \Delta d / d_o = \left(\frac{1+\nu}{E}\right)\sigma_{\phi}\sin^2\psi + \left(\frac{\nu}{E}\right)(\sigma_{11}+\sigma_{22}) \qquad (1)$$

gdzie: d_o – odległość między płaszczyznami sieciowymi w materiale nie odkształconym ($d_o^{311}{}_{(aust)}$=0.1083 nm), σ_{11} i σ_{22} - naprężenia główne w płaszczyźnie powierzchni materiału (ze względu na głębokość pomiaru nie przekraczającej kilkunastu μm przyjmuje się σ_{33}=0), ν - współczynnik Poissona, E - moduł Younga [16].

Naprężenia obliczono przyjmując rentgenowskie stałe sprężystości dla płaszczyzn (311) ½ s_2=6.33×10^{-6} MPa^{-1} i $-s_1$=1.42×10^{-6} MPa^{-1} (program XRDWin), których wartości odpowiadają stałym mechanicznym modułu Younga E=200 GPa i współczynnika Poissona ν=0.29, według zależności (2).

$$E = 1/(s_1 + \tfrac{1}{2}s_2) \quad \text{i} \quad \nu = -s_1/(s_1 + \tfrac{1}{2}s_2)\,. \qquad (2)$$

3. Wyniki

Wyznaczone w pomiarach wartości naprężeń na powierzchni zewnętrznej rury, według schematu na rysunku 3b, przedstawiono na rysunkach 5 i 6. Wszystkie wyznaczone naprężenia były rozciągające, zarówno w kierunku obwodowym, jak i wzdłużnym. Wartości naprężeń cechowały się dużym rozrzutem, większym z przypadku naprężeń wzdłużnych – zakres 21-253 MPa, w porównaniu do naprężeń w kierunku obwodowym – zakres 65-227 MPa.

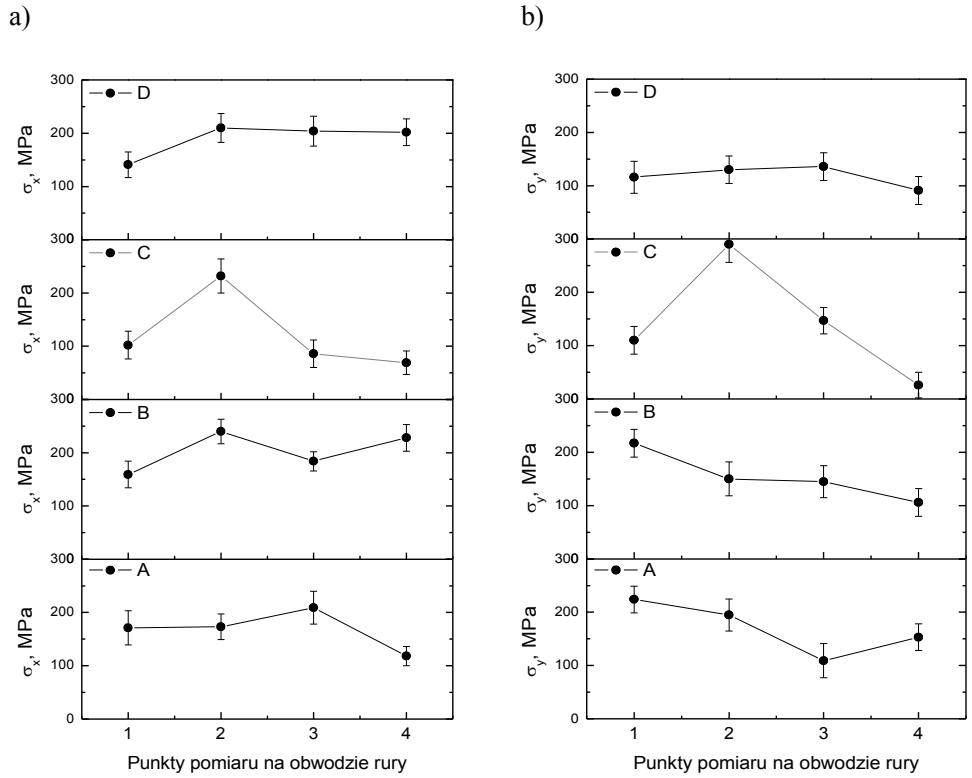


Fig. 5. Rozłożenie naprężeń a) obwodowych i b) wzdłużnych na obwodzie rury

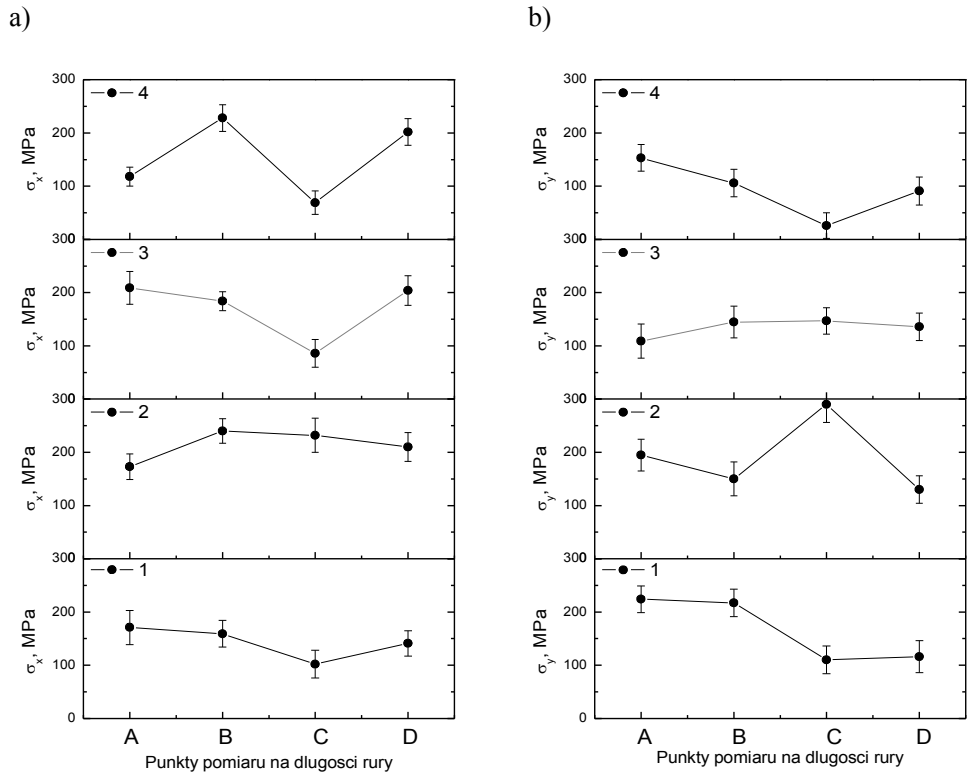


Fig. 6. Rozłożenie naprężeń a) obwodowych i b) wzdłużnych na długości rury

Analiza rozłożenia naprężeń na obwodzie rury (Rys. 5) wykazała, że największe, co do wartości, zróżnicowanie naprężeń w obu badanych kierunkach występowało na obwodzie C, w centralnej części długości rury. Najbardziej jednorodne rozłożenie naprężeń stwierdzono na obwodzie D, z tym, że naprężenia obwodowe były średnio większe od naprężeń wzdłużnych o ok. 80 MPa.

Analiza rozłożenia naprężeń na długości rury (Rys. 6) wykazała, że największe co do wartości zróżnicowanie naprężeń obwodowych występowało na tworzących 4 i 3, a naprężeń wzdłużnych na tworzącej 2. Najbardziej jednorodne rozłożenie naprężeń obwodowych stwierdzono na tworzących 1 i 2, a naprężeń wzdłużnych na tworzącej 3.

Bardziej uogólnione informacje o rozłożeniu naprężeń szczątkowych na powierzchni rury mogą dostarczyć uśrednione wartości naprężeń dla poszczególnych obszarów rury (Rys. 7) oraz współczynniki zmienności (Tabela 1) definiowane jako (3)

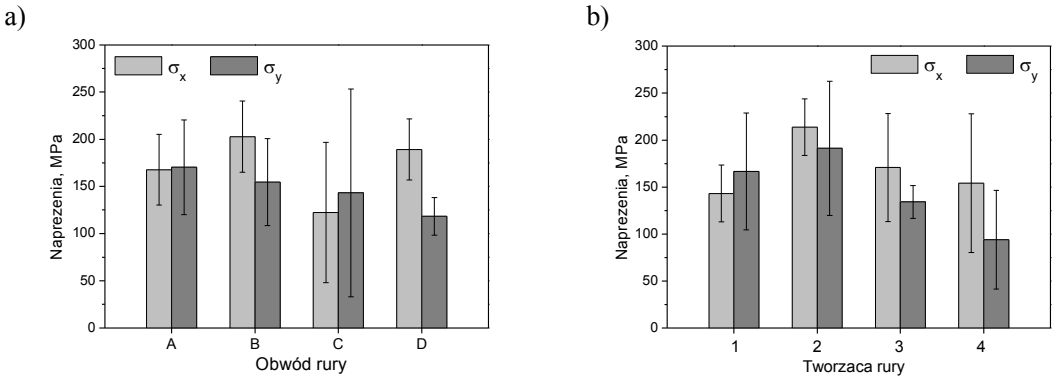
$$V = S/\overline{\sigma} \tag{3}$$

gdzie: S – odchylenie standardowe, $\overline{\sigma}$ - średnia arytmetyczna.

W analizie średnich wartości naprężeń $\overline{\sigma}_x$ i $\overline{\sigma}_y$ na poszczególnych obwodach rury (A, B, C i D) nie wykazano między nimi żadnej zależności (Rys. 7a). Odnotować można tylko, że w centralnej części rury (na obwodzie C) naprężenia obwodowe $\overline{\sigma}_x$ były najmniejsze. Średnie naprężenia w kierunku wzdłużnym $\overline{\sigma}_y$ (Rys. 7a) były największe w pobliżu końca rury reprezentowanego obwodem A, i malały sukcesywnie w kierunku obwodu D. Dodatkowo stwierdzono, że zarówno naprężenia obwodowe, jak i naprężenia wzdłużne, wyznaczone w obszarze C cechowały największe współczynniki zmienności V (61% i 77%), co oznacza duże rozproszenie wartości naprężeń na tym obwodzie rury. Z kolei w obszarze D, naprężenia w obu kierunkach cechowały najmniejsze współczynniki zmienności (17%), czyli występowało małe rozproszenie wartości naprężeń.

W analizie średnich wartości naprężeń $\overline{\sigma}_x$ i $\overline{\sigma}_y$ na poszczególnych tworzących rury (1, 2, 3 i 4) (Rys. 7b) wykazano podobną tendencję w ich rozłożeniu. Największe średnie naprężenia charakteryzowały tworzącą 2, a najmniejsze tworzącą 4. Te tworzące znajdowały się na przeciwległych ściankach rury. Duże rozproszenie wartości naprężeń w obu kierunkach występowało wzdłuż tworzącej 4 (V= 48% i 56%). Małe rozproszenie wykazały naprężenia obwodowe na tworzącej 2, a naprężenia wzdłużne na tworzącej 3.

Mimo podobieństwa w rozłożeniu naprężeń obwodowych i wzdłużnych na tworzących rury należy podkreślić, że analiza rozkładu naprężeń wzdłużnych nie jest często postrzegana jako szczególnie użyteczna w przypadku rur. Wynika to z faktu, że pod wpływem ciśnienia medium wewnątrz eksploatowanej rury zwiększeniu ulegają przede wszystkim naprężenia obwodowe. Ich wartości na powierzchni zewnętrznej są dwukrotnie większe od naprężeń wzdłużnych (osiowych). Z tego powodu ryzyko uszkodzeń rur związane jest przede wszystkim z wartościami naprężeń obwodowych - i wyznaczenie tych naprężeń zlecają producenci i odbiorcy. W niniejszej pracy badano rurę w stanie technologicznym. Pod wpływem ciśnienia cieczy podczas hydroformowania, nie dochodzi do swobodnego poszerzania się rury ze względu na ograniczenie jej kształtu matrycą, stąd można przypuszczać, że naprężenia wzdłużne mogą być relatywnie większe w stosunku do obwodowych. Wyniki eksperymentalnie wyznaczonych naprężeń potwierdzają słuszność tego przypuszczenia w odniesieniu do powierzchni zewnętrznej rury.

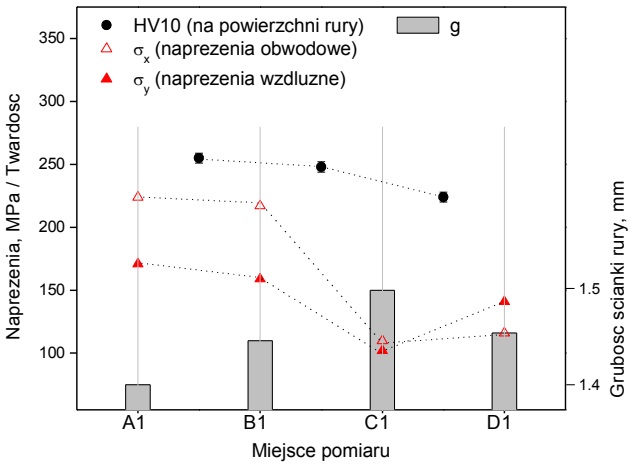


Rys. 7. Uśrednione wartości naprężeń w poszczególnych obszarach hydroformowanej rury a) na obwodach i b) na tworzących (znaczniki błędu przedstawiają odchylenie standardowe

Tabela 1. Współczynniki zmienności naprężeń *V* w różnych obszarach rury hydroformowanej

Obwód	$V(\sigma_x)$	$V(\sigma_y)$	Tworząca	$V(\sigma_x)$	$V(\sigma_y)$
A	22	29	1	21	37
B	19	30	2	14	37
C	61	77	3	34	13
D	17	17	4	48	56

Ze względu na kształt rury oraz niewielką grubość ścianki nie było możliwe wykonanie dokładnych pomiarów twardości bezpośrednio na powierzchni rury, ani metodą Vickersa, ani metodą ultradźwiękową. Niemniej, na rysunku 8 przedstawiono wyniki próby takich pomiarów wykonane w przybliżeniu wzdłuż tworzącej 1. Pomiary wykonano na powierzchni rury pomiędzy obwodami A, B, C i D, celem nie uszkodzenia miejsc wytypowanych do pomiaru naprężeń. Mimo tylko poglądowego charakteru wyznaczonych twardości można zauważyć, że odzwierciedliły one poniekąd rozłożenie naprężeń σ_x i σ_y wzdłuż tej tworzącej, tzn. były mniejsze tam gdzie naprężenia również były mniejsze. Po pomiarach naprężeń z rury pobrano wycinki celem wyznaczenia grubości ścianki, które przedstawiono również na rysunku 8. Potwierdzono przypuszczenie, że naprężenia są największe w obszarach rury z najmniejszą grubością ścianki, zredukowanej podczas kształtowania.



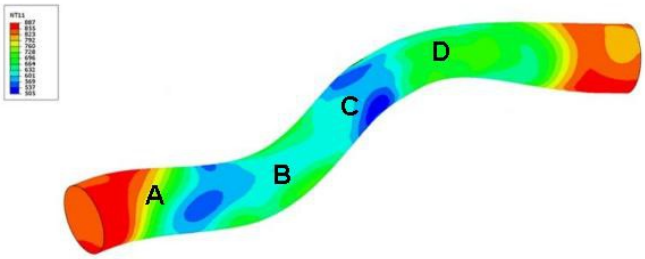
Rys. 8. Porównanie naprężeń, twardości i grubości ścianki na długości rury (tworząca 1)

4. Dyskusja i podsumowanie

Na powierzchni zewnętrznej rury do układu wydechowego, wykonanej ze stali chromowo-niklowej kształtowanej technologią hydroformowania, stwierdzono złożony stan naprężeń rozciągających. Wartości największych naprężeń powierzchniowych przekraczają poziom granicy plastyczności stali 304L w wersji DDQ (ok. 170MPa). Oznacza to, że stal podczas hydroformowania uległa umocnieniu, i jak można przypuszczać mając na względzie złożony kształt rury, również wskutek wstępnego gięcia poprzedzającego hydroformowanie.

Naprężenia cechowały się szerokim zakresem zmienności zarówno w kierunku obwodowym (σ_x), jak i w kierunku wzdłużnym (σ_y), wynoszącym odpowiednio 69-240 MPa i 26-290 MPa. W analizie średnich wartości naprężeń oraz współczynników zmienności wykazano, że najmniejsze naprężenia i o jednocześnie największym rozproszeniu wartości występowały w centralnej części rury, gdzie redukcja grubości ścianki była najmniejsza. Z uwagi na powierzchniowy charakter wyznaczonych naprężeń źródła tego rozproszenia należy poszukiwać w odmiennych warunkach tarcia różnych fragmentów powierzchni rury o matrycę podczas hydroformowania.

Rozłożenie naprężeń powierzchniowych wyznaczonych metodą dyfraktometryczną ogólnie współgra z modelem rozłożenia odkształceń w poszczególnych obszarach rury wygenerowanym komputerowo, przedstawionym na rysunku 9. Należy jednak podkreślić, że przy grubościach ścianki jakie posiadała badana rura (<1.5 mm) model prezentuje raczej odkształcenia średnie w całej grubości ścianki. Nie odzwierciedli on ewentualnych, incydentalnych zjawisk (np. związanych z transportem i magazynowaniem), które mogą wystąpić w warunkach produkcji.



Rys. 9. Rozłożenie odkształceń w hydroformowanej rurze wygenerowane w modelowaniu komputerowym

Wartości wyznaczonych naprężeń i analiza ich rozłożenia oparte na pomiarach metodą rentgenowską dotyczą niewielkich powierzchniowych obszarów, określonych przekrojem skolimowanego promieniowania (∅2mm) oraz jego wnikaniem w stal (ok. 17 μm) przy powierzchni zewnętrznej rury. Zobrazowanie rozłożenia naprężeń, czy odkształceń, w tak cienkiej warstwie nie jest osiągalne w komputerowym modelowaniu procesów kształtowania plastycznego. Stąd, metoda rentgenowska może służyć jako cenne uzupełnienie modelowania, szczególnie że nowoczesne dyfraktometry umożliwiają dokonanie nieniszczących pomiarów na cienkościennych wyrobach o złożonym kształcie powierzchni.

Z prezentowanych w literaturze analiz numerycznych przepływu ciekłego medium w przewodach rurowych silnika wynika, że wywiera ono różne ciśnienie na ścianki przewodu w różnych jego miejscach. Ciśnienie to zależy głównie od kątów zagięcia przewodu oraz parametrów medium, takich jak temperatura, gęstość lub szybkość przepływu [20]. Przemawia to za tym, aby w miejscach, gdzie krzywizna przewodu rurowego jest najsilniejsza dokonywać mapowania rozłożenia naprężeń poprzez kompleksowe pomiary - na obwodach i wzdłuż tworzących rury.

Literatura

[1] Alaswad A, Benyounis K Y, Olabi A G. Tube hydroforming process: a reference guide. Materials and Design 2012; 33: 328-339.

[2] Baczmański A, Wierzbanowski K, Lipiński P. Determination of Residual Stresses in Plastically Deformed Polycrystalline Material. Materials Science Forum 1994; 157-162: 2051-2058.

[3] Bahman K, Trends for stainless steel tube in automotive applications. The Tube & Pipe Journal, September 13, 2005 (thefabricator.com).

[4] Brytan Z, Stainless steel in the automotive industry (in Polish). STAL Metale & Nowe Technologie 2013; 11-12: 14-19.

[5] Chałupczak J. Hydromechanical spreading in application to the formation of tees and X-pieces (in Polish). Works of the Kielce University of Technology. Mechanics; 39. Habilitation dissertation. Kielce, 1986.

[6] Gronostajski Z, Kuziak R. Metallurgical, technological and functional foundations of advanced high-strength steels for the automotive industry (in Polish). Works of the Institute of Ferrous Metallurgy 2010; 22-26.

[7] Hashemi R, Assemoiur A, Masourni E, Abad K. Implementation of the forming limit stress diagram to obtain suitable load path in tube hydroforming considering M–K model. Materials & Design 2009; 30(9): 3545-3553.

[8] ISSF International Stainless Steels Forum. Stainless Steel Consumption Forecast, October 2017, (http://www.worldstainless.org/statistics) 05.12.2018

[9] Kocańda A, Sadłowska H. Automotive component development by means of hydroforming. Archives of Civil and Mechanical Engineering 2008; 8(3): 55-69.

[10] Koç M. An overall review of tube hydroforming (THF) technology. Journal of Materials Processing Technology 2001; 108: 384-393.

[11] Koç M (Ed.). Hydroforming for Advanced Manufacturing. Woodhead Publishing Limited England, and CRC Press USA, 2008.

[12] Kucharska B, Krzywiecki M. Stresses in a Cr-Ni superficial steel layer based on x-ray measurements and electropolishing Solid State Phenomena 2015; 223: 348-354.

[13] Kucharska B., Wróbel A., Kulej E., Nitkiewicz Z. The X-ray measurement of the thermal expansibility of Al-Si alloy in the form of cast and a protective coating on steel. Solid State Phenomena 2010; 163: 286-290.

[14] Miłek T. Variations of wall thickness in the sections of hydromechanically bulged copper cross joints. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2003; 2(18): 45-48.

[15] Morphy G. Pressure-sequence and high pressure hydroforming: Knowing the processes can mean boosting profits. The Tube & Pipe Journal, September/October 1998 (thefabricator.com, February 2001).

[16] Skrzypek S J, Witkowska M, Kowalska M, Chruściel K. The non-destructive X-Ray methods in measuring of some material properties (in Polish). Hutnik-Wiadomości Hutnicze 2012; 79(4): 238-246.

[17] Susceptibility of stainless steels to plastic working. Euro Inox, Series: Materials and applications. Book No. 8, 2008.

[18] Wróbel-Knysak A, Kucharska B, The abrasion of Al-Si coatings with different silicon crystal morphology used in car exhaust systems. Tribologia 2016; 5: 209-218.

[19] Xianfeng Chen, Zhongqi Yu, Bo Hou, Shuhui Li, Zhomgqin Lin. A theoretical and experimental study on forming limit diagram for a seamed tube hydroforming. Journal of Materials Processing Technology 2011; 211(12): 2012-2021.

[20] Kumbár V, Votava J, Numerical modelling of pressure and velocity rates of flowing engine oils in real pipe. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2015; 7(3): 422-426.

dr inż. Mariusz Kosobudzki

Katedra Inżynierii Maszyn Roboczych i Pojazdów Przemysłowych
Politechnika Wrocławska
ul. Łukasiewicza 7/9, 50-371 Wrocław
E-mail: mariusz.kosobudzki@pwr.edu.pl

dr inż. Mariusz Stańco

Katedra Konstrukcji i Badań Maszyn
Politechnika Wrocławska
ul. Łukasiewicza 7/9, 50-371 Wrocław
E-mail: mariusz.stanco@pwr.edu.pl

Problemy oceny trwałości wybranego elementu pojazdu na podstawie przyspieszonego testu przebiegowego

Słowa kluczowe: *resor, stabilizator, test przebiegowy, badania przyspieszone, trwałość, uogólniony wskaźnik trwałości, ciężarowy samochód terenowy*

Streszczenie: W artykule przedstawiono wyniki analizy trwałości elementów sprężystych występujących w zawieszeniu specjalnego terenowego pojazdu ciężarowego 4x4 wykorzystując dane uzyskane podczas przyspieszonego testu drogowego przeprowadzonego podczas jazdy off-road. Wskazano na występujące ograniczenia w dostępie do danych materiałowych jakie są obecne na etapie wstępnego doboru podzespołu (brak danych wytrzymałości zmęczeniowej) oraz podano alternatywną analityczną metodę szacowania wytrzymałości zmęczeniowej. Wskazano na powstające różnice w uzyskanych wynikach oraz na najważniejsze ich źródła. Przedstawiono również sposób wykorzystania uogólnionego wskaźnika trwałości *d* jako parametru niezależnego od danych materiałowych podzespołu, który można wykorzystać do oceny wpływu obciążeń wynikowych (rejestrowanych) powstających podczas ruchu pojazdu w ustalonych warunkach drogowych na trwałość analizowanego podzespołu i powiązać ich wartość z rodzajem testowego odcinka drogowego.

1. Wstęp

Ocena trwałości pojazdu jest zagadnieniem bardzo złożonym i trudnym do przeprowadzenia, a jednocześnie niezbędnym w procesie osiągania gotowości konstrukcji pojazdu do uruchomienia jego produkcji seryjnej [15]. W przypadku obiektów złożonych, np. specjalnych terenowych pojazdów ciężarowych, od których wymaga się dużej niezawodności połączonej z wymaganą trwałością, proces projektowo – konstrukcyjny organizuje się według odpowiedniego modelu zarządzania. Przykładem takiego modelu może być opracowany przez NASA model V [12]. W modelu tym przyjmuje się, że przejście do kolejnego etapu procesu projektowo – konstrukcyjnego jest możliwe wtedy, kiedy poprzedni etap zostanie oceniony pozytywnie. Do oceny etapu konieczne jest przeprowadzenie odpowiednich testów, których stopień złożoności i pracochłonności jest uzależniony od oceny stopnia wpływu danego etapu na jakość wyrobu końcowego. Taką ocenę można przeprowadzić wykorzystując np. metodę Design for Six Sigma [17], która pozwala wskazać, z jaką dokładnością należy oceniać poszczególne etapy procesu projektowo – konstrukcyjnego, żeby produkt końcowy charakteryzował się wymaganą trwałością,

czy niezawodnością. Kluczowe zatem jest wybranie i przeprowadzenie odpowiednich testów oddających z wystarczającą dokładnością wpływ obciążeń przewidywanych do planowanych warunków eksploatacji na trwałość wyrobu [11,12,14].

W przypadku specjalnych terenowych pojazdów ciężarowych problem doboru odpowiednich testów wydaje się szczególnie trudny. Są to pojazdy produkowane w małych seriach, przeznaczone są do jazdy w zmiennych warunkach drogowych ze zmiennym obciążeniem przez długi okres czasu eksploatacji (do 30 lat). Konieczne i kluczowe staje się zatem przyjęcie szeregu założeń upraszczających dotyczących między innymi miejsca i warunków prowadzenia badań.

Niektórzy producenci pojazdów swoje testy prowadzą na sparametryzowanych drogowych odcinkach pomiarowych, np. Tatra [22], które powinny być reprezentatywne do rzeczywistych warunków drogowych, na których określa się stopień wpływy profilu drogi i parametrów ruchu pojazdu na wartość powstających obciążeń, a w efekcie końcowym na trwałość analizowanych podzespołów. Badania odbiorcze gotowego pojazdu prowadzi w imieniu przyszłego użytkownika wskazana jednostka certyfikująca, realizując ustalony program badań. Wynikiem przeprowadzonych badań jest wydanie lub odmowa wydania certyfikatu o zgodności wyrobu z wymaganiami odbiorcy. Jednak efekty tych badań są dostępne dopiero wtedy, kiedy pojazd jest gotowy do produkcji.

Są również tacy producenci pojazdów, którzy nie dysponują dostępem do ośrodków badań drogowych, co istotnie utrudnia prowadzenie testów. W takim przypadku testy prowadzi się na wybranych dostępnych odcinkach drogowych, w tym na drogach publicznych. Pojawia się jednak problem skorelowania obciążeń przyjmowanych jako reprezentatywne (występujące na odcinkach testowych jednostki certyfikującej) z tymi wykorzystywanymi przez producenta pojazdu. Stąd producenci pojazdów poszukują różnych parametrów, które można wykorzystać do porównania warunków badań występujących w jednostce certyfikującej z warunkami badań własnych.

Ze względu na ograniczony czas oraz zasoby finansowe, ale również np. niedostatek danych, badania zasadniczo prowadzi się do osiągnięcia stanu granicznego badanego podzespołu lub do momentu, kiedy na podstawie zgromadzonych danych będzie można określić zależność pomiędzy warunkami użytkowania a trwałością podzespołu. Do powiązania powstających obciążeń (warunków ruchu) i trwałości elementu wykorzystuje się różne modele przebiegu degradacji. Są wśród nich również takie, których wykorzystanie nie wymaga znajomości szczegółowych danych materiałowych uzyskiwanych na drodze eksperymentalnych badań stanowiskowych, które są bardzo czasochłonne i kosztowne.

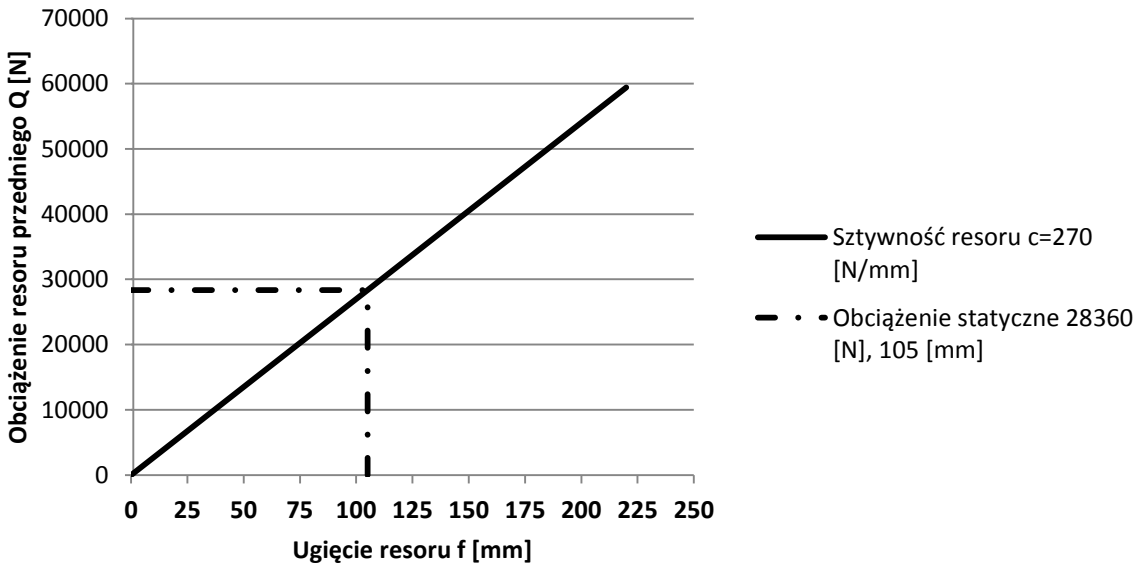
Uzyskane wyniki trwałości badanych podzespołów często odnosi się do wielkości opisujących wykorzystany zasób pracy, np. w jednostkach przebiegu pojazdu (km), godzin pracy (mtg) i innych zgodnie z wymaganiami przyszłego użytkownika. Otrzymane wyniki, przy przyjęciu założeń upraszczających, np. że drogowe odcinki testowe i ustalone parametry ruchu są stałe, pozwalają powiązać przebieg jednostkowy pojazdu ze stopniem jego degradacji. W efekcie otrzymuje się dane pozwalające porównywać ze sobą wpływ wybranych rodzajów drogowych odcinków testowych (tych w jednostce certyfikującej z odcinkami własnymi) i ustalonych parametrów ruchu pojazdu na stopień degradacji wybranego podzespołu. Przykłady szacowania trwałości podzespołów pojazdów można znaleźć w literaturze, np. [2,5,10,20]. Problemem jednak dalej pozostaje wskazanie parametru, którego wyznaczenie mogłoby być wykorzystane jako wskaźnik porównawczy do wstępnego oszacowania trwałości podzespołu w powiązaniu z wybranym drogowym odcinkiem testowym.

2. Cel i zakres badań

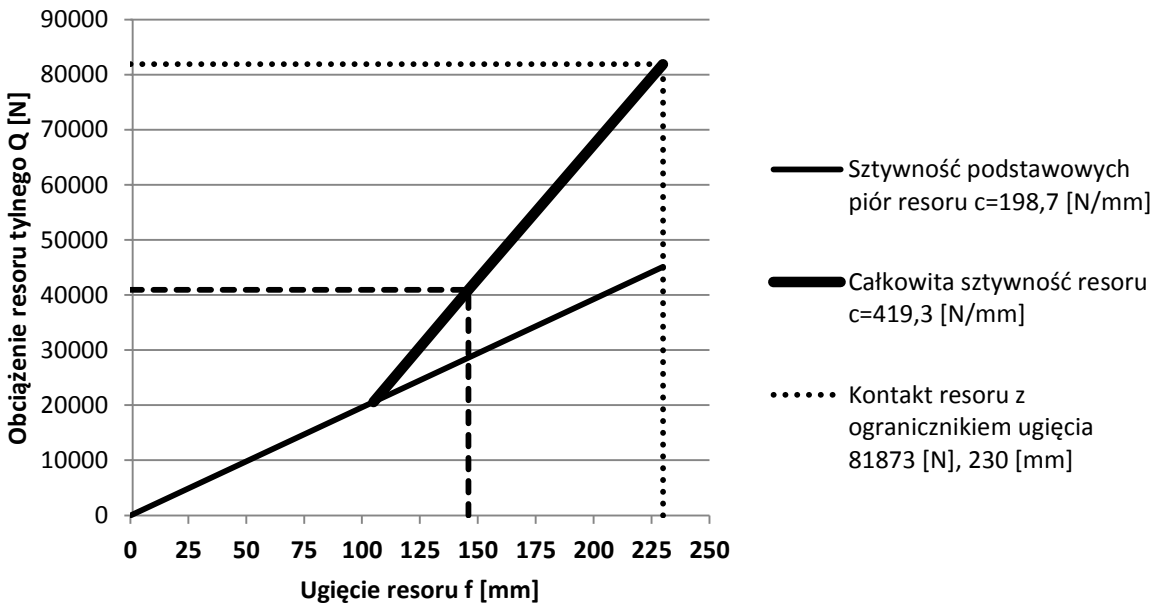
Celem badań było oszacowanie trwałości wybranych podzespołów terenowego pojazdu ciężarowego w wybranych warunkach ruchu (off-road) oraz sprawdzenie, czy można do wstępnej oceny trafności doboru tych podzespołów do pojazdu zastosować uogólniony wskaźnik trwałości, do wyznaczenia którego nie jest wymagana pełna wiedza dotycząca wytrzymałości materiału z którego zostały wykonane podzespoły, co jest typowym problemem występującym podczas przyspieszonych badań przebiegowych. Wybranie uogólnionego wskaźnika trwałości i wyznaczenie jego wartości w warunkach badań odbiorczych w jednostce certyfikującej pozwoliłoby na ustalenie ekwiwalentnego programu badań w oparciu o dostępne dla producenta odcinki drogowe. Szczegółowa charakterystyka wykorzystanego w badaniach uogólnionego wskaźnika trwałości nie została przedstawiona w niniejszej pracy, ale została opisana w publikacji [6].

Przedmiotem przyspieszonych badań przebiegowych były elementy sprężyste (resory paraboliczne, stabilizatory) występujące w zawieszeniu specjalnego terenowego pojazdu ciężarowego 4x4, przeprowadzone w warunkach ruchu off-road przy ograniczonych danych dotyczących wytrzymałości materiału z którego te elementy zostały wykonane. Ograniczenie badań do jednego rodzaju odcinka drogowego zostało określone przez producenta pojazdu.

W zawieszeniu analizowanego pojazdu elementami podatnymi były resory paraboliczne, które pozwalają na ruch względny kół i nadwozia w osi pionowej, a jednocześnie odbierają swobodę ruchu w pozostałych osiach oraz stabilizatory, które swoim działaniem zmniejszają przechył boczny nadwozia poprawiając w ten sposób stateczność ruchu pojazdu [16]. Wymienione podzespoły pracują w złożonym stanie naprężenia, ale często w celu uproszczenia badań przyjmuje się, że resory poddawane są zginaniu, a drążki stabilizatora skręcaniu [1]. Na rysunkach 1 i 2 przedstawiono charakterystyki sztywności resorów oraz zakresy ugięcia przy różnym obciążeniu pojazdu.



Rys. 1. Charakterystyka sztywności resoru parabolicznego osi przedniej



Rys. 2. Charakterystyka sztywności resoru parabolicznego osi tylnej

Stabilizatory zostały wykonane z prętów o przekroju kołowym. Podstawowe wymiary charakterystyczne przedstawiono w tabeli 1.

Tabela 1. Zestawienie podstawowych wymiarów charakterystycznych stabilizatorów

	Stabilizator osi przedniej	**Stabilizator osi tylnej**
Długość części poddawanej skręcaniu [mm]	730	820
Ramię działania siły wywołującej skręcanie [mm]	520	340
Średnica części poddawanej skręcaniu [mm]	40	50
Wskaźnik wytrzymałości przekroju na skręcanie [cm^3]	6,28	12,27

Materiałem wykorzystanym do produkcji wymienionych podzespołów była stal 51CRV4 (R_m=1350 MPa). Z deklaracji producenta wynikało, że pióra resorów zostały poddane obróbce cieplnej, a twardość rdzenia pióra powinna wynosić zgodnie z normą [15] od 363 do 460 HB. Dodatkowo, pióra resorów po stronie rozciąganej poddane zostały procesowi kulowania. Dzięki takiemu zabiegowi na tej powierzchni wprowadzone zostały naprężenia normalne ściskające, które znacznie redukują wartości naprężeń rozciągających powstających podczas pracy elementu [3,18]. Ze względu na brak danych o wartościach tych naprężeń oraz o głębokości ich wprowadzenia w strukturę materiału, do ich oszacowania wykorzystano dostępne dane, przedstawione np. [9,13,15]. Na podstawie tych danych przyjęto, że w stanie nieobciążonym naprężenia ściskające mogą osiągać wartość od 300 do 400 MPa, a głębokość wprowadzonych naprężeń może wynosić 15 - 25 μm.

3. Model wytrzymałości zmęczeniowej analizowanych podzespołów

Badania trwałości podzespołów ograniczonej zmęczeniem materiału, z którego zostały wykonane są przedsięwzięciem złożonym i czasochłonnym. Do badań eksperymentalnych wybiera się liczbę próbek, której liczność zależy od etapu procesu projektowo –

konstrukcyjnego, liczby analizowanych poziomów obciążenia oraz ilości powtórzeń testu. Na wstępnym etapie doboru podzespołu wystarcza zazwyczaj liczba próbek od 6 do 12, która wzrasta do 24 przy testach niezawodnościowych [1]. Liczbę powtórzeń testu można wyznaczyć z zależności [10]:

$$LP = 100(1 - \frac{SL}{n_s}) \quad (1)$$

gdzie: LP – procentowa liczba powtórzeń, SL – liczba poziomów obciążenia, n_s – liczba próbek.

Na etapie testów wstępnych procentowa liczba powtórzeń wynosi od 17 – 33. W tabeli 2 zestawiono liczbę próbek, poziomy obciążenia i liczbę powtórzeń odpowiednio dla 12 i 24 próbek.

Tabela 2. Zestawienie liczby próbek, poziomów obciążenia oraz procentowej liczby powtórzeń

Liczba próbek n_s	**Poziomy obciążenia SL**	**Procentowa liczba powtórzeń LP**
12	2	83,3
12	3	75,0
24	3	87,5
24	4	83,3
24	5	79,2
24	6	75,0

Przedstawione dane obrazują czasochłonność badań eksperymentalnych podzespołu w celu określenia jego charakterystyki wytrzymałości zmęczeniowej. W zrealizowanych badaniach, które miały charakter testów wstępnych gotowych podzespołów sprawdzonych przez producenta ograniczono się do oceny poprawności ich doboru do pojazdu. Badania ograniczono do jednego pojazdu. Ze względu na brak szczegółowych danych dotyczących wytrzymałości zmęczeniowej (wyznaczonej eksperymentalnie krzywej S-N) konieczne było jej wyznaczenie na drodze obliczeń teoretycznych i powiązanie otrzymanych wyników z parametrem łączącym trwałość podzespołu z rodzajem drogowego odcinka testowego [6].

Do obliczenia wytrzymałości zmęczeniowej resoru wykorzystano zależności, które pozwalały wyznaczyć wykres zmęczenia w oparciu o ograniczony zbiór danych. Przyjęto, że kluczowe jest wyznaczenie wytrzymałości zmęczeniowej w zakresie wysokocyklowym, tzn. w granicach od 10^3 do 10^6 liczby cykli. Sposób wyznaczania poszczególnych wartości zaczerpnięto z dostępnych publikacji, m.in. [4,10].

Wytrzymałość zmęczeniową dla 10^3 liczby cykli wyznaczono na podstawie zależności

$$A_{NC,R} = A_{1000} \cdot C_R \quad (2)$$

gdzie: $A_{NC,R}$ – amplituda naprężeń dla obciążeń niskocyklowych z uwzględnieniem współczynnika niezawodności C_R, A_{1000} – amplituda naprężeń dla obciążeń niskocyklowych, C_R – współczynnik niezawodności.

Wartość A_{1000} można wyznaczyć z zależności:

$$A_{1000} = \alpha_{NC} \cdot R_m \quad (3)$$

gdzie: R_m – granica wytrzymałości materiału wyznaczona w próbie statycznego rozciągania, α_{NC} – współczynnik zależny od typu obciążenia dla 10^3 cykli; 0,9 dla zginania, 0,72 dla skręcania.

Wartość współczynnika niezawodności C_R zależy od oczekiwanej pewności działania podzespołu. W badaniach przyjęto wstępnie, że C_R=1.

Wytrzymałość zmęczeniowa dla 10^6 liczby cykli wyznaczono na podstawie zależności, w której uwzględniono wskaźniki korygujące:

$$A_{WC,R} = A_{WC} \cdot C_L \cdot C_S \cdot C_D \cdot C_R \quad (4)$$

gdzie: $A_{WC,R}$ – amplituda naprężeń dla obciążeń wysokocyklowych z uwzględnieniem współczynnika niezawodności C_R, A_{WC} – amplituda naprężeń dla obciążeń wysokocyklowych, C_L – współczynnik typu obciążenia, C_S – współczynnik stanu powierzchni, C_D – współczynnik zależny od wielkości elementu, C_R – współczynnik niezawodności.
Wartość A_{WC} można wyznaczyć z zależności:

$$A_{WC} = \alpha_{WC} \cdot R_m \tag{5}$$

gdzie: R_m – granica wytrzymałości materiału wyznaczona w próbie statycznego rozciągania, α_{WC} – współczynnik zależny od rodzaju materiału dla 10^6 cykli; dla stali (R_m<1400 MPa) wynosi 0,5.

Wartość współczynnika typu obciążenia C_L przyjęto na podstawie danych dostępnych w literaturze [10]. Dla zginania C_L=1, a dla skręcania C_L=0,58.

Wartość współczynnika stanu powierzchni C_S można określić na podstawie pomiaru chropowatości powierzchni i wartości R_m materiału. Dostarczane przez producenta podzespoły były fabrycznie zabezpieczone przed szkodliwym działaniem warunków atmosferycznych farbą ochronną. Pomiar rzeczywistej chropowatości powierzchni wiązałby się ze skutecznym usunięciem tej warstwy. Ze względu na istniejące ograniczenia nie dokonano pomiaru chropowatości, a do wyznaczenia wartości współczynnika C_S wykorzystano dostępne dane z literatury [10], w których uwzględniono rodzaj obróbki plastycznej, jakiej poddano badane podzespoły. Resory są poddawane walcowaniu i kulowaniu, co wprowadza w strukturę materiału naprężenia ściskające, częściowo kompensujące powstające podczas pracy podzespołu naprężenia rozciągające. Przyjęto do obliczeń wartość współczynnika C_S równą 0,76.

Wartość współczynnika zależnego od wielkości elementu C_D obliczono na podstawie zależności [10]:

$$C_D = 1{,}189 \cdot d^{-0{,}097} \tag{6}$$

gdzie: d – średnica elementu, mm.
Dla elementu o przekroju prostokątnym (pióra resoru), średnicę zastępczą można wyznaczyć z zależności [10]:

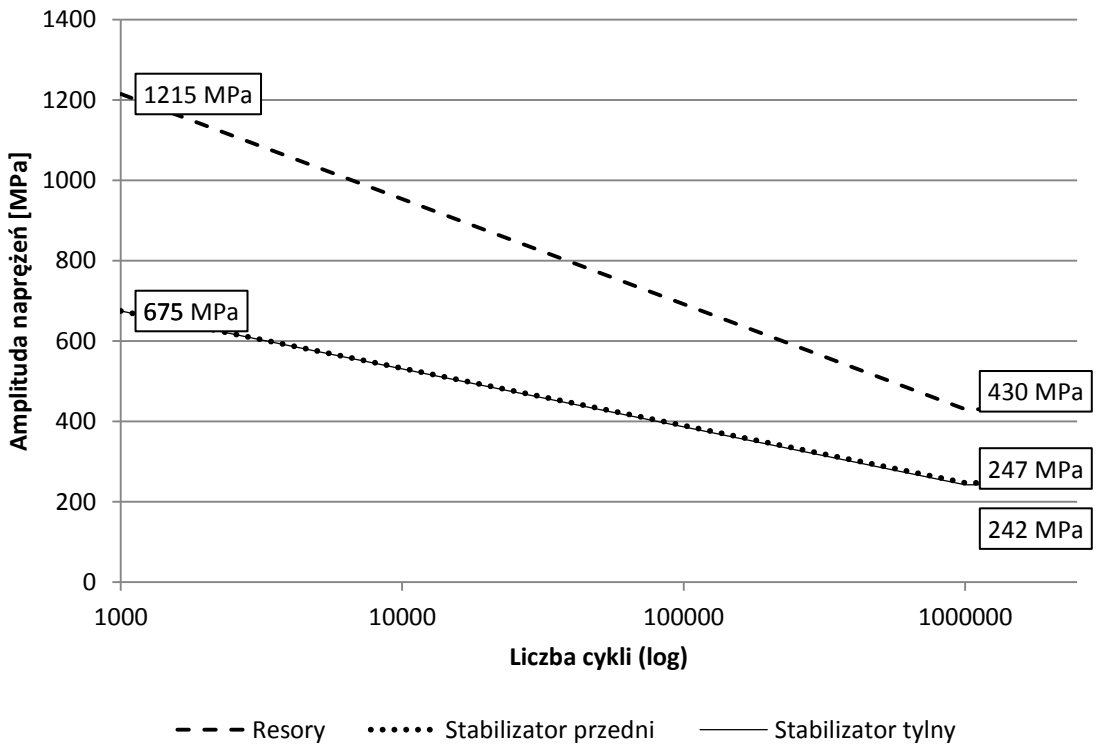
$$d_z = \sqrt{0{,}65 \cdot s \cdot w} \tag{7}$$

gdzie: s – szerokość przekroju, w – wysokość przekroju.
Obliczone wartość C_D przedstawiono w tabeli 3.

Tabela 3. Wartości współczynnika C_D wyznaczone do analizowanych elementów

	Resor przedni	Resor tylny	Stabilizator przedni	Stabilizator tylny
C_D	0,84	0,84	0,83	0,81

Na podstawie wyznaczonych danych sporządzono wykresy wytrzymałości zmęczeniowej resorów i stabilizatorów, co zostało przedstawione na rysunku 3. Wyznaczone przebiegi wytrzymałości zmęczeniowej stabilizatora przedniego i tylnego są bardzo zbliżone do siebie, a różnica występująca w okolicach nieograniczonej wytrzymałości zmęczeniowej jest niewielka i wynosi 5 MPa. Wyznaczone do tego obszaru wartości wytrzymałości wynoszą odpowiednio: stabilizator przedni 247 MPa, a stabilizator tylny 242 MPa.



Rys. 3. Wyznaczone wykresy wytrzymałości zmęczeniowej resorów i stabilizatorów

Otrzymane wykresy wytrzymałości zmęczeniowej podzespołów wykorzystano do analizy trwałości badanych podzespołów.

4. Przebieg badań

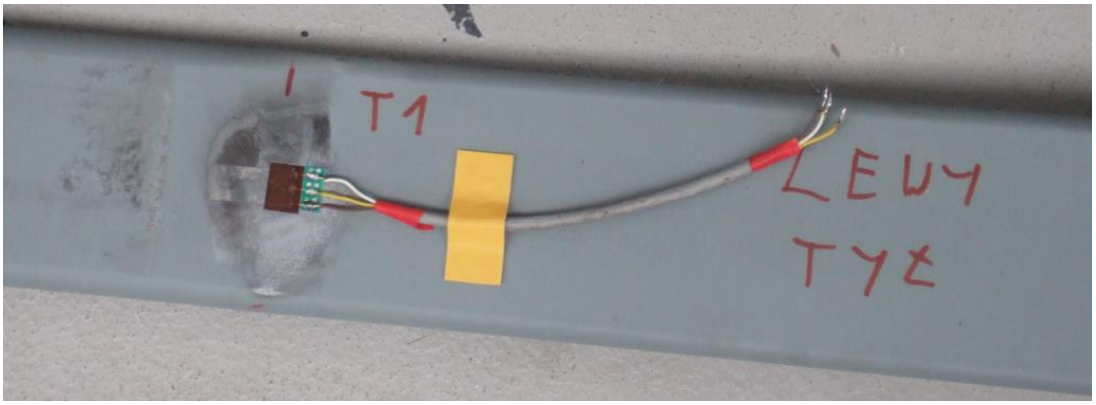
Badania zostały przeprowadzone w warunkach poligonowych na terenie Akademii Wojsk Lądowych we Wrocławiu. Wybrany odcinek piaszczystej drogi off-road stanowił pętlę pomiarową o długości około 1 km. Ze względu na charakter nierówności średnia prędkość jazdy wynosiła ok. 7 km/h. Była ona ustalona na podstawie wcześniejszych przejazdów i wniosków z dotychczasowych badań [7,8]. Wybrany drogowy odcinek pomiarowy odpowiadał warunkom poligonowym, które są brane pod uwagę przy projektowaniu pojazdu do oczekiwanych warunków ruchu opisanych w profilu eksploatacji pojazdu [11]. Nie był to jednak odcinek sparametryzowany. Pojazdem kierował kierowca testowy producenta. Pojazd do badań został obciążony równomiernie wykorzystując całkowitą ładowność.

Resory pojazdów samochodowych pracują w złożonym stanie naprężenia [1,18]. W przyspieszonych badaniach przebiegowych jest jednak trudno rejestrować wszystkie pojawiające się obciążenia i oceniać ich wpływ na trwałość zmęczeniową resoru. Stąd przyjmuje się, że dominującym obciążeniem jest zginanie, które wywołuje naprężenia normalne w przekrojach piór resorów. W przypadku stabilizatorów, są one zaprojektowane w taki sposób, aby ulegały skręcaniu. Przyjęcie przedstawionych uproszczeń powoduje, że gromadzenie danych koniecznych do dalszych analiz sprowadza się do rejestracji pojawiających się naprężeń wywołanych zginaniem resorów i skręcaniem stabilizatorów. Osiągnięta w ten sposób redukcja danych jest krokiem przemyślanym, wynikającym z ekonomiki czasu i dostępnych zasobów oraz ograniczonych danych o analizowanych podzespołach. W tabeli 4 przedstawiono zbiór charakterystyk, jakimi dysponowano na etapie wstępnego doboru podzespołów.

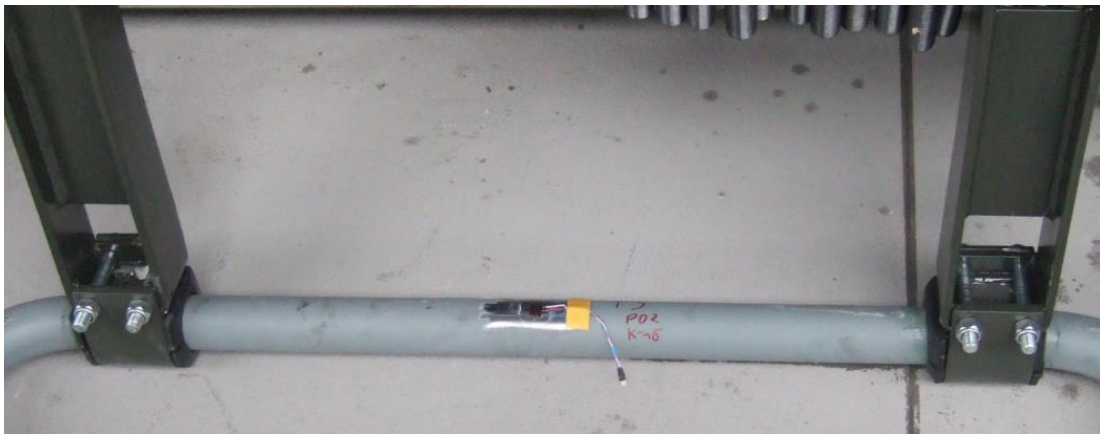
Tabela 4. Podstawowe dane o analizowanych podzespołach

Resory	charakterystyka sztywności, wymiary, masa, materiał, rodzaj obróbki cieplnej i plastycznej, deklarowana twardość na powierzchni, deklarowana trwałość,
Stabilizatory	charakterystyka sztywności, wymiary, masa, materiał, rodzaj obróbki cieplnej i plastycznej,

Układ pomiarowy wykorzystany w przyspieszonych badaniach przebiegowych elementów sprężystych pojazdu składał się z czujników tensometrycznych naklejonych na przygotowanych powierzchniach piór resorów i stabilizatorach (rys. 4÷5). Czujniki tensometryczne zostały naklejone w miejscach, gdzie spodziewano się uzyskać największe wartości naprężenia (okolice jarzma mocującego pióra resorów, a w przypadku stabilizatora w połowie długości odcinka ulegającego skręcaniu). Wybór miejsc został dodatkowo potwierdzony w oparciu o posiadany model MES resorów [19,20], co nie jest krokiem standardowym.



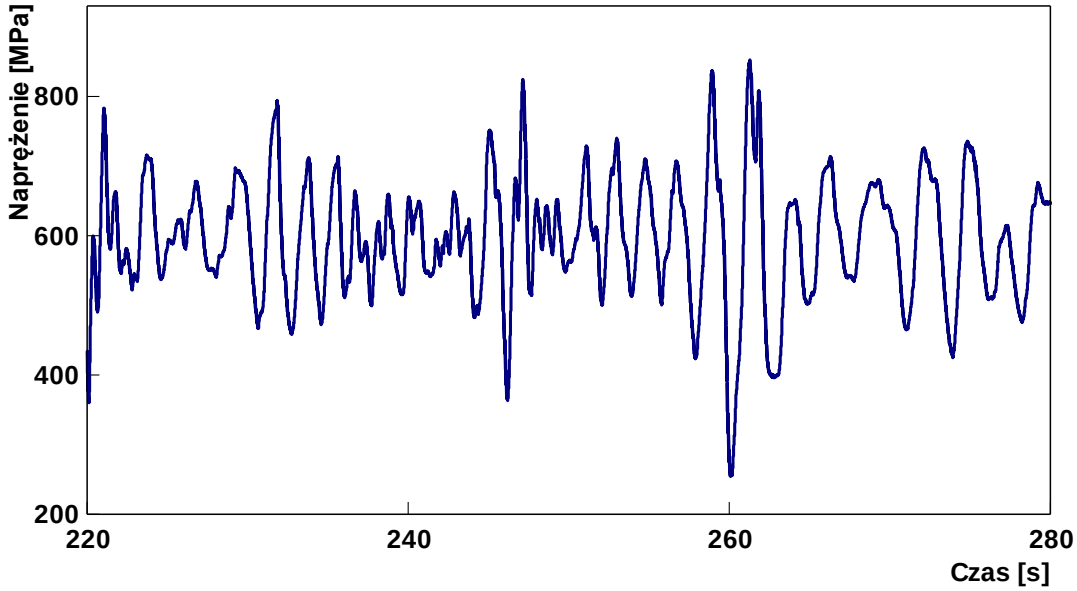
Rys. 4. Pióro resoru z naklejonym tensometrem



Rys. 5. Stabilizator z naklejonym tensometrem

Podczas prób drogowych rejestrowano przebiegi obciążenia, które następnie poddano filtracji Rainflow wyznaczając cykle obciążenia. Przykładowy przebieg obciążenia pióra resoru osi tylnej przedstawiono na rysunku 6. Filtracja Rainflow polegała na wyznaczeniu i zliczeniu cykli obciążenia z zarejestrowanego przebiegu. Metoda ta jest dzisiaj powszechnie stosowana i podlega standaryzacji. Resory po zamontowaniu do pojazdu ulegają początkowemu obciążeniu masą własną pojazdu i ładunku, co wpływa na niesymetryczność obciążeń powstających podczas uginania i odginania tych elementów podczas jazdy (przesunięcie wartości średniej). Do uwzględnienia tego efektu wykorzystano model

Goodmana [21]. Do sumowania uszkodzeń zmęczeniowych wykorzystano hipotezę Palmgrena-Minera, która zakłada liniowe kumulowanie się uszkodzeń aż do osiągnięcia wartości granicznej, przyjmowanej jako 1. Jest to model powszechnie wykorzystywany w obliczeniach zmęczeniowych.



Rys. 6. Przykładowy przebieg naprężeń pióra resoru zainstalowanego do osi tylnej pojazdu (wartości nie uwzględniają wstępnych naprężeń wywołanych kulowaniem)

5. Analiza otrzymanych wyników

Na podstawie zarejestrowanych przebiegów i wyznaczonej teoretycznie wytrzymałości zmęczeniowej oszacowano trwałość badanych podzespołów, którą podano w jednostkach przebiegu pojazdu. Otrzymane wyniki są wstępną informacją wykorzystywaną do oceny trafności wyboru podzespołów do pojazdu przy założeniu, że występujące podczas badań obciążenia są reprezentatywne do przyszłych przewidywanych warunków eksploatacji. W zestawieniu wyników zebranych w tabeli 5 można zauważyć istotny rozrzut otrzymanych wartości do poszczególnych podzespołów. Przyczyną tego rozrzutu jest brak dokładnych danych dotyczących rzeczywistej wartości naprężeń wstępnych wprowadzonych do piór resorów, które należało oszacować.

Tabela 5. Zestawienie prognozowanej trwałości elementów dla różnych wartości naprężeń wstępnych

Wielkość	**Resor LP**	**Resor PP**	**Resor LT**	**Resor PT**	**Stab. przód**	**Stab. tył**
Zasięg (bez uwzględnienia naprężeń wstępnych ściskających)[km]	44	35	12	10	555	600
Zasięg (naprężenia wstępne ściskające 300 MPa) [km]	56698	45708	13878	11338	-	-
Zasięg (naprężenia wstępne ściskające 350 MPa) [km]	121229	100819	30231	24357	-	-
Zasięg (naprężenia wstępne ściskające 400 MPa) [km]	273596	236196	71713	56939	-	-

Przedstawione w tabeli 5 dane wskazują, że na obliczoną trwałość resoru bardzo duży wpływ ma poprawnie przyjęta wartość naprężeń wstępnych ściskających, możliwych do

określenia na podstawie np. oceny głębokości zmian mikrostruktury materiału, będącej efektem kulowania. Taką ocenę można przeprowadzić np. wykonując badania materiałowe niszczące podzespołu [9,13]. Przekazywana od producenta informacja ogólna o przeprowadzonej obróbce plastycznej, bez szczegółowych danych, jest niewystarczająca do poprawnego skalkulowania trwałości podzespołu.

W tabeli 6 przedstawiono, jaki wpływ na obliczoną trwałość podzespołów ma redukcja obciążeń, na które pośrednio lub bezpośrednio ma wpływ styl jazdy kierowcy. Z przedstawionych danych wynika, że redukcja obciążeń o 5% (np. zmniejszenie prędkości jazdy, zmiana trasy przejazdu, regulacja ciśnienia powietrza w ogumieniu, itp.) może wydłużyć trwałość podzespołu o ok. 50%, a redukcja obciążeń o 10%, zwiększa trwałość o ok. 100%.

Tabela 6. Wpływ wartości obciążeń na trwałość podzespołu

Podzespół	**Trwałość przy obciążeniu zarejestrowanym (bez uwzględnienia wstępnych naprężeń ściskających)**	**Trwałość przy obciążeniu zredukowanym o 5%**	**Trwałość przy obciążeniu zredukowanym o 10%**
Resor przedni prawy	35	48	68
Resor tylny prawy	10	15	21
Stabilizator przedni	600	776	1016
Stabilizator tylny	555	718	941

Na podstawie przedstawionych danych można zauważyć, że próba wyznaczenia trwałości podzespołu ograniczonej wytrzymałością zmęczeniową w przyspieszonym teście przebiegowym nastręcza szereg trudności i może być obarczona istotnym błędem, wynikającym np. z przyjęcia przybliżonych wartości wielkości pośrednich. Jako istotne ograniczenia w wyznaczeniu dokładnych wartości można wskazać m. in. brak danych dotyczących eksperymentalnie wyznaczonej wytrzymałości zmęczeniowej podzespołu, co zmusza do przeprowadzenia przybliżonych obliczeń teoretycznych, brak szczegółowych danych materiałowych podzespołu (rzeczywista wartość R_m, wartość wprowadzonych naprężeń ściskających i ich głębokości) oraz parametrów opisujących stan warstwy wierzchniej (chropowatość). Ponadto, we wstępnych badaniach przebiegowych, w sytuacji braku dostępu do sparametryzowanych torów testowych pojawia się potrzeba porównywania efektów zastosowania nowych rozwiązań konstrukcyjnych podzespołów w odniesieniu do tych wcześniej wykorzystywanych i oceny ich pracy w powiązaniu z rodzajem drogowego odcinka testowego wykorzystywanego w jednostce certyfikującej. Parametrem użytecznym w rozwiązaniu tego typu problemu może być wielkość nazywana uogólnionym wskaźnikiem trwałości d [6], która wyraża w postaci liczbowej całościowy wpływ parametrów opisujących ruch pojazdu (m. in. prędkość, rodzaj odcinka testowego) na trwałość podzespołu, ale bez odniesień do charakterystyki materiałowej podzespołu.

Koncepcja wykorzystania uogólnionego wskaźnika trwałości d została opisana w [6] i opiera się na wyznaczeniu wartości wyrażenia:

$$d = \sum n_i A_i^{\beta} \tag{8}$$

gdzie: d – uogólniony wskaźnik trwałości (pseudo damage), A_i – amplituda obciążenia wyznaczona np. metodą Rainflow, n_i – liczba cykli obciążenia o amplitudzie A_i, β – współczynnik nachylenia krzywej zmęczeniowej (można przyjąć wstępnie, że dla elementów wykonywanych bez specjalnych zabiegów wykańczających powierzchnię materiału (np. szlifowanie, polerowanie) współczynnik β=5).

Przedstawiony uogólniony współczynnik trwałości *d* wykorzystano do przedstawienia różnić w obciążeniach takich samych podzespołów odpowiednio dla lewej i prawej strony. Przykładowe wyniki przeprowadzonych obliczeń przedstawiono w tabeli 7.

Tabela 7. Zestawienie wartości uogólnionego wskaźnika trwałości dla resorów przednich i tylnych odpowiednio dla lewej i prawej strony pojazdu.

Wartość uogólnionego wskaźnika trwałości *d*	**Resor przedni lewy**	**Resor przedni prawy**	**Resor tylny lewy**	**Resor tylny prawy**
$d_{100\%}$	$6{,}74*10^{16}$	$8{,}48*10^{16}$	$2{,}42*10^{17}$	$2{,}80*10^{17}$
$d_{95\%}$	$4{,}71*10^{16}$	$6{,}08*10^{16}$	$1{,}72*10^{17}$	$1{,}99*10^{17}$
$d_{90\%}$	$3{,}33*10^{16}$	$4{,}30*10^{16}$	$1{,}18*10^{17}$	$1{,}40*10^{17}$
d_{300MPa}	$5{,}19*10^{13}$	$6{,}43*10^{13}$	$2{,}12*10^{14}$	$2{,}59*10^{14}$
d_{350MPa}	$2{,}43*10^{13}$	$2{,}92*10^{13}$	$9{,}73*10^{13}$	$1{,}21*10^{14}$
d_{400MPa}	$1{,}07*10^{13}$	$1{,}24*10^{13}$	$4{,}10*10^{13}$	$5{,}19*10^{13}$

Przedstawione w tabeli 7 wartości uogólnionego wskaźnika trwałości *d* dotyczą przypadków, kiedy zostały uwzględnione wartości naprężeń zmierzonych ($d_{100\%}$), wartości naprężeń zredukowanych o 5% i 10% ($d_{95\%}$, $d_{90\%}$) oraz wartości uwzględniające wstępne naprężenia ściskające wynoszące odpowiednio 300 MPa, 350 MPa i 400 MPa (d_{300MPa}, d_{350MPa}, d_{400MPa}). Rosnąca wartość parametru *d* wskazuje na bardziej destrukcyjny przebieg obciążeń. Z danych przedstawionych w tabeli 7 wynika, że resor prawy przedni, który jest taki sam jak lewy przedni, podczas badań podlegał bardziej destrukcyjnym obciążeniom. Podobnie resor prawy tylny był bardziej obciążony zmęczeniowo niż resor tylny lewy. Widać również, że trwałość elementów sprężystych w zawieszeniu osi przedniej jest wyższa niż elementów w osi tylnej. Otrzymane wartości uogólnionego wskaźnika trwałości podzespołu *d* nie reprezentują jednak rzeczywistej trwałości podzespołu, wyrażają tylko w postaci liczbowej (łatwej do porównywania), czy w danych warunkach ruchu obciążenia działające na podzespół są bardziej lub mniej destrukcyjne w porównaniu z innym takim samym podzespołem.

6. Podsumowanie

Podstawowym celem badań było zidentyfikowanie obciążeń działających na elementy sprężyste zawieszenia oraz oszacowanie ich trwałości ograniczonej wytrzymałością zmęczeniową, co zostało przedstawione w tabeli 5. i wykorzystane jako wstępne dane do oceny trafności doboru tych elementów do pojazdu. Dodatkowym celem było wskazanie parametru, którego wykorzystanie pozwalałoby ocenić, w jakim stopniu warunki ruchu i rodzaj drogowego odcinka pomiarowego wpływają na wartość obciążeń wybranych elementów, ograniczając ich trwałość.

Trwałość analizowanych podzespołów jest kluczowa do oszacowania niezawodności pojazdu, rozumianego jako system techniczny, którego obciążenia wynikające z warunków ruchu zmieniają się w szerokim zakresie (od jazdy po drogach twardych bez ładunku po jazdę off-road z ładunkiem). Przedstawione analizy opierały się na ograniczonych danych dostępnych na etapie wstępnego doboru nowego podzespołu do pojazdu. Otrzymane wyniki trwałości podzespołów przedstawiono w przeliczeniu na teoretyczny zasięg jazdy pojazdu, który jest skutecznym parametrem porównawczym. Ze względu na ograniczone dane, do obliczeń przyjęto wariantowo różne wartości obciążeń (jako wynik możliwych zmian

w stylu jazdy kierowcy) oraz wstępnych naprężeń ściskających pióra resorów wykazując, jaki wpływ mają one na analizowany przebieg pojazdu.

Badania zostały ograniczone tylko do jednego pojazdu (jednego kompletu analizowanych podzespołów) poruszającego się z ustaloną prędkością w wybranych warunkach drogowych, stąd otrzymane wyniki stanowią tylko materiał wstępny do dalszych analiz. Warto jednak zauważyć, że wykorzystanie proponowanego uogólnionego wskaźnika trwałości *d* daje możliwość wstępnego porównania trwałości poszczególnych resorów pojazdu. Rozróżnienie stopnia degradacji takich samych resorów, ale różnie obciążonych (co wynika z różnego ukształtowania podłoża pod każdym z kół podczas przejazdu) wskazuje, że można również porównywać wartości parametru *d* wyznaczone dla tego samego podzespołu (tutaj resoru) w różnych warunkach drogowych (różne odcinki testowe). Jeżeli dodatkowo przeprowadzi się normalizację parametru *d* i sprowadzi jego wartość do jednostkowej długości odcinka pomiarowego (np. do 1 km), będzie można szacować stopień degradacji tego samego podzespołu w różnych warunkach ruchu i na różnych odcinkach testowych. To daje powody by przypuszczać, że można odtworzyć skutek działania obciążeń zarejestrowanych na jednym odcinku testowym (np. w jednostce certyfikującej) za pomocą innego dostępnego odcinka testowego (dostępnego u producenta pojazdu), co byłoby innowacyjnym wykorzystaniem parametru *d* wyznaczonego na podstawie przekształconego równania Basquina. Potwierdzenie tego przypuszczenia będzie jednak wymagało dodatkowych badań.

Literatura

1. ASTM E739-10 – Standard practice for statistical analysis of linear or linearized stress-life (S-N) and strain-life (ε-N) fatigue data. USA: ASTM International, West Conshohocken, PA, 2015.
2. Johannesson P, Speckert M. (editors) Guide to load analysis for durability in vehicle engineering. London: John Willey & Sons, 2014.
3. Hryciów Z, Krasoń W, Wysocki J. The experimental tests of the friction coefficient between the leaves of the multi-leaf spring considering a condition of the friction surfaces. Eksploatacja i Niezawodność – Maintenance and Reliability 2018; 20(4): 682-688. http://dx.doi.org/10.17531/ein.2018.4.19.
4. Kocańda S, Szala J. Podstawy obliczeń zmęczeniowych. Warszawa: PWN, 1991.
5. Kosobudzki M. Metoda szacowania trwałości ustroju nośnego pojazdu wysokiej mobilności. Rozprawa doktorska. Politechnika Wrocławska, 2013.
6. Kosobudzki M, Smolnicki T. Generalized vehicle durability index for different traffic conditions, AIP Conference Proceedings 2019; 2078: 020017(1-6); http://doi.org/10.1063/1.5092020.
7. Kosobudzki M, Stanco M. The experimental identyfication of torsional angle on a load-carrying truck frame during static and dynamic tests. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2016; 18(2): 285–290. http://dx.doi.org/10.17531/ein.2016.2.17.
8. Kosobudzki M. The use of acceleration signal in modeling process of loading an element of underframe of high mobility wheeled vehicle. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16(4): 595-599.
9. Kukiełka L, Bartosik P, Szyc M. Optymalizacja procesu kulowania strumieniowego ze względu na naprężenia wynikowe, Archiwum Technologii Maszyn i Automatyzacji, KBM PAN_Oddział w Poznaniu 2010; 30/1: 117-126.
10. Lee Y-L, Pan J, Hathaway R, Barkey M. Fatigue testing and analysis. Theory and practice. Elsevier, 2005.

11. MIL-STD_810G – Environmental engineering considerations and laboratory tests. USA: Department of Defense test method standard, 2008.
12. NASA System Engineering Handbook. USA: NASA Headquarters, Washington, 2007.
13. Nasiłowska B, Bogdanowicz Z, Brzeziński M, Mońka G, Zasada D. Wpływ kulowania na strukturę, mikrotwardość i naprężenia własne stali austenitycznej 1.4539. Biuletyn WAT 2015; 64/2: 103-110.
14. Norma Obronna NO-06-A101 – Uzbrojenie i sprzęt wojskowy. Ogólne wymagania techniczne, metody kontroli i badań. Postanowienia ogólne
15. Norma PN-90/S-47250 – Pojazdy samochodowe i przyczepy. Resory piórowe. Wymagania i badania.
16. Reimpell J, Betzler J. Podwozia samochodów. Podstawy konstrukcji. Warszawa: Wydawnictwo Komunikacji i Łączności WKŁ, 2008.
17. Rusiński E, Koziołek S, Jamroziak K. Quality assurances metod for the design and manufacturing process of armoured vehicles. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2009; 43(3): 70-77.
18. Spring design manual. AE-21. USA: SAE International, 1996.
19. Stańco M, Iluk A. Numeryczno – doświadczalna analiza wytężenia resoru parabolicznego pojazdu ciężarowego. Materiały konferencyjne XVI Konferencji Naukowo – Technicznej TKI2016 – Techniki Komputerowe w Inżynierii, 18-21.10.2016.
20. Stańco M. Analysis of the influence of leaf geometry on stiffness and effort of the heavy-duty spring. In: Rusiński E, Pietrusiak D. (editors) Proceedings of the 14th International Scientific Conference - Computer Aided Engineering. Springer International Publishing, 2018;
21. Łagoda T, Macha E. Trwałość zmęczeniowa maszyn – laboratorium. Opole: Politechnika Opolska, 2005.
22. www.tatratrucks.com/your-tatra-partner/tatra-testing-grounds/

Dr hab. inż. Zbigniew Kamiński, prof. nzw.

Bialystok University of Technology Faculty of Mechanical Engineering, Wiejska 45C, 15-351 Białystok, Poland
Email: z.kaminski@pb.edu.pl

Mgr inż. Paweł Radzajewski

Bialystok University of Technology Faculty of Mechanical Engineering, Wiejska 45C, 15-351 Białystok, Poland
Email: radzajewskilech@wp.pl

OBLICZENIA OPTYMALNEGO ROZDZIAŁU SIŁ HAMUJĄCYCH W PRZYCZEPACH ROLNICZYCH KATEGORII R3 i R4

CALCULATIONS OF THE OPTIMAL DISTRIBUTION OF BRAKE FORCE IN AGRICULTURAL VEHICLES CATEGORIES R3 AND R4

Keywords: *Brake force distribution, optimization, agricultural vehicles, braking systems*

Keywords: *Rozkład siły hamowania, optymalizacja, pojazdy rolnicze, układy hamulcowe*

Abstract
Fulfilling the requirements of the EU Directive 2015/68 in the area of braking for agricultural trailers depends on the proper selection of individual components of the braking system. This paper describes the requirements regarding braking performance and distribution of brake forces in agricultural trailers in R3 and R4 categories. On this basis, a methodology for calculating the optimal linear distribution of brake forces, characteristic for agricultural trailers with pneumatic braking systems, has been developed. The examples of calculation of an optimal distribution of brake forces for a two- and three-axle trailer with a tandem suspension system of the rear axle assembly have been provided. The optimization algorithm with the Monte Carlo method has been described, based on which a computer program was developed to select a linear distribution of brake forces in a three-axle trailer with 'walking beam' and 'bogie' suspensions. The presented calculations can be used in the design process to select the parameters of wheel braking mechanisms and then the characteristics of the pneumatic valves of the braking system.

Abstract
Spełnienie wymagań Dyrektywy UE 2015/68 w zakresie hamowania przyczep rolniczych zależy od właściwego doboru poszczególnych komponentów układu hamulcowego. W pracy opisano wymagania dotyczące skuteczności hamowania oraz rozdziału sił hamujących w przyczepach rolniczych kategorii R3 i R4. Na tej podstawie opracowano metodykę obliczeń optymalnego liniowego rozdziału sił hamujących, charakterystycznego dla przyczep rolniczych z pneumatycznymi układami hamulcowymi. Zamieszczono przykłady obliczeń optymalnego rozdziału sił hamujących dla przyczepy dwu i trzyosiowej z tandemowym układem zawieszenia zespołu osi tylnych. Opisano algorytm optymalizacji metodą Monte Carlo, na podstawie którego opracowano program komputerowy do doboru liniowego rozdziału sił hamujących w przyczepie trzyosiowej z zawieszeniem „walking beam" i „bogie". Przedstawione obliczenia można wykorzystać w procesie projektowania do doboru parametrów kołowych mechanizmów hamulcowych, a następnie charakterystyk zaworów pneumatycznych układu hamulcowego.

1. Wprowadzenie

W przyczepach i holowanych maszynach rolniczych stosowane są najczęściej pneumatyczne lub hydrauliczne układy hamulcowe zasilane i sterowane z ciągnika rolniczego [4, 8, 17, 26, 27, 28]. Obecnie inercyjne hamulce najazdowe można stosować tylko w wolnobieżnych pojazdach ciągniętych ($v \leq 40$ km/h) o masie całkowitej mniejszej niż 8000 kg i w pojazdach szybkobieżnych ($v > 40$ km/h), których masa całkowita nie przekracza 3500 kg [2]. Do uruchomienia hamulców zasadniczych traktora wykorzystuje się napęd mechaniczny, hydrauliczny lub powietrzny. Wybór rodzaju napędu i źródła energii zależy od konstrukcji i masy ciągnika.

W ciągnikach małej i średniej mocy stosuje się proste i niedrogie hydrauliczne układy hamulcowe bez wspomagania [14]. W ciągnikach większej mocy stosuje się przede wszystkim układy hydrauliczne wspomagane z układu hydrauliki siłowej traktora oraz układy pneumatyczne [17, 19, 27]. W ciągnikach małej mocy nadal atrakcyjne ze względu na koszty są hamulce uruchamiane mechaniczne.

Współdziałanie układu hamulcowego ciągnika i układu hamulcowego przyczepy zapewnia montowany w ciągniku zawór sterujący hamulcami przyczepy (pneumatyczny lub hydrauliczny). W zależności od rodzaju zastosowanych hamulców roboczych ciągnika stosuje się zawory sterujące hamulcami przyczepy uruchamiane mechanicznie, pneumatycznie lub hydraulicznie [12, 13, 22].

Wysokowydajne układy hamulcowe są krytyczną cechą współczesnych pojazdów rolniczych. Obowiązujące od 2016 r. nowe rozporządzenie UE w sprawie pojazdów rolniczych [2] obejmuje szereg nowatorskich i znacznie wyższych wymagań w zakresie skuteczności hamowania ciągników i przyczep, kompatybilności, standardów bezpieczeństwa i stabilności, a także tym wprowadzenie układów ABS dla pojazdów poruszających się z prędkością ponad 60 km/h.

Dla wszystkich kategorii pojazdów ciągniętych zwiększono wymaganą wartość wskaźnika hamowania, a dla pojazdów o masie całkowitej ponad 3500 kg (przyczepy rolnicze kategorii R3 i R4 oraz holowane maszyny rolniczych kategorii S2) i poruszających się z prędkością ponad 40 km/h wprowadzono wymóg określonego rozdziału sił hamujących między osie pojazdu. Dzięki temu można spełnić wymóg osiągnięcia dostatecznie dużego opóźnienia względnego (wskaźnika hamowania, czyli ilorazu opóźnienia do przyspieszenia ziemskiego $z=d/g$), warunkujący osiągnięcie krótkiej drogi hamowania oraz zapewnić stateczność kursową hamowanego pojazdu w każdych warunkach ruchu. Podobnie jak w przepisach dotyczących pojazdów samochodowych [24] nie sformułowano oddzielnych zaleceń dla zestawów drogowych, traktując poszczególne człony zestawu tak, jakby były pojedynczymi pojazdami.

W celu dopasowania rozdziału sił hamowania między ciągnikiem a pojazdem ciągniętym wprowadzono po raz pierwszy wymogi dotyczące kompatybilności w postaci dopuszczalnych obszarów zmian wskaźników hamowania pojazdu ciągnącego i ciągniętego w funkcji ciśnienia w przewodzie sterującym. Spełnienie wymogów kompatybilności, jak i wymogów dużej szybkości działania przy hamowaniu nagłym (czas reakcji mniejszy równy 0,6 s [2]), sprzyja skróceniu drogi hamowania zestawów ciągnikowych i zmniejszeniu sił w sprzęgu [21].

Przyjęcie nowego ustawodawstwa stawia wysokie wymagania producentom pojazdów rolniczych pod względem układów hamulcowych [5]. Łączne spełnienie wymagań dotyczących skuteczności, stateczności i kompatybilności hamowania przyczep rolniczych zależy od właściwego doboru i obliczeń poszczególnych komponentów układu hamulcowego (mechanizmów i siłowników hamulcowych, zaworów i regulatorów sił hamowania) z uwzględnieniem parametrów konstrukcyjnych przyczep, w tym układu osi i rodzaju zastosowanego zawieszenia [1, 11, 16].

Obliczenia inżynierskie układów hamulcowych przyczep rolniczych dzieli się na projektowe (synteza) i sprawdzające (analiza). Celem obliczeń projektowych jest wyznaczanie podstawowych parametrów konstrukcyjnych układów hamulcowych i ich elementów z uwzględnieniem zadanych charakterystyk roboczych. Obliczenia projektowe obejmują między innymi:
- wyznaczenie dopuszczalnego rozdziału sił hamujących,
- dobór charakterystyk korektorów sił hamowania,
- obliczenia sił i momentów hamujących kół poszczególnych osi dla zadanego rozdziału sił hamujących,
- obliczenia lub dobór mechanizmów hamulcowych,
- dobór aktuatorów i obliczenia mechanizmów uruchamiania,
- wybór koncepcji układu hamulcowego i dobór jego elementów (zaworów, przewodów itp.).

Obliczenia sprawdzające mają na celu zbudowanie i analizę charakterystyk rozpatrywanego układu, gdy znane są jego parametry konstrukcyjne. Obliczenia sprawdzające obejmują między innymi:
- obliczenie skuteczności hamowania przy zadanym poziomie minimalnego ciśnienia w układzie hamulcowym,
- obliczenie charakterystyk statycznych, w tym przebiegów wskaźników przyczepności wykorzystanej przez poszczególne osie w funkcji ciśnienia w siłowniku hamulcowym (wskaźnika hamowania) oraz sprawdzenie przebiegów wskaźników hamowania ciągnika i przyczepy w funkcji ciśnienia na złączu sterującym (pasm kompatybilności układu hamulcowego),
- obliczenia charakterystyk dynamicznych w celu sprawdzenia szybkości działania (czas reakcji) i synchronii działania poszczególnych obwodów układu hamulcowego.

W niniejszej pracy opisano metodykę optymalnego doboru rozdziału sił hamujących w przyczepach rolniczych kategorii R3 i R4. Obliczenia rozdziału sił hamujących stanowią podstawę obliczeń projektowych układów hamulcowych pojazdów, gdyż mają znaczący wpływ na dobór podstawowych mechanizmów i elementów układu hamulcowego oraz osiągi układu hamulcowego [23]. Zamieszczono przykład obliczeń liniowego rozdziału sił hamujących w przyczepie dwuosiowej oraz przyczepie trzyosiowej z tandemowych układem zawieszenia zespołu osi tylnych. Opisano algorytm optymalizacji liniowego rozdziału sił hamujących w przyczepie trzyosiowej z zastosowaniem klasycznej metody Monte Carlo.

2. Wymagania dotyczące skuteczności, stateczności i kompatybilności układów hamulcowych pojazdów kategorii R3, R4 i S2

Dobierając rozdział sił hamujących pomiędzy osie przyczepy (maszyny holowanej) należy dążyć do rozkładu idealnego. Wówczas wskaźniki f_i przyczepności wykorzystanej przez wszystkie osie są jednakowe podczas całego procesu hamowania, a tym samym równe wskaźnikowi z hamowania pojazdu:

$$\frac{T_1}{R_1} = \frac{T_2}{R_2} = \ldots = \frac{T_i}{R_i} = f_i = z \qquad (1)$$

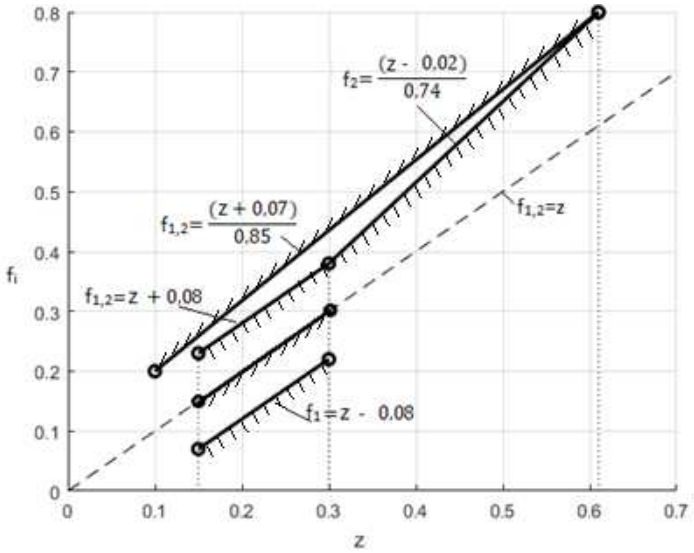
gdzie: T_i – siła hamowania kół i-tej osi, R_i – reakcje pionowe nawierzchni drogi na koła i-tej osi, z – wskaźnik hamowania pojazdu $z=\sum T_i/\sum R_i$.

Taki rozdział sił hamowania przyjęto uważać za optymalny, gdyż na nawierzchni homogenicznej uzyskuje się największą możliwą w danych warunkach intensywność hamowania i spełnienie z zapasem wymogów dotyczących skuteczności hamowania (tabela 1).

Tabela 1. Wymagana skuteczność hamowania hamulców roboczych przyczep rolniczych [2]

Kategoria pojazdu	Wskaźnik hamowania z [%] przy p=6,5 bar	
	v≥30 km/h	v>30 km/h
Przyczepy R2, R3, R4	35%	50%
Holowane maszyny S2	35%	50%

Ze względu na zmienny stopień załadowania przyczep osiągnięcie idealnego rozdziału sił hamujących jest praktycznie niemożliwe, nawet przy zastosowaniu regulatorów sił hamujących. Dlatego dla szybkobieżnych pojazdów rolniczych (prędkość powyżej 40 km/h) wyznaczono dopuszczalne granice odstępstwa wskaźników wykorzystania przyczepności f_i poszczególnych osi od rozdziału idealnego. Od 2016 roku dopuszcza się dwa rozwiązania pokazane na rys.1 [2].



Rys.1. Graniczne wartości wykorzystania przyczepności dla obu rozwiązań

Pierwsze rozwiązanie: wskaźnik wykorzystania przyczepności dla każdej z osi musi spełniać warunek zapewnienia minimalnej wymaganej skuteczności hamowania:

$$\left.\begin{matrix} f_1 \\ f_2 \end{matrix}\right\} \le \frac{z+0.07}{0.85} \quad \text{dla} \quad z = 0{,}1 \div 0{,}61 \tag{2}$$

oraz warunek wcześniejszego blokowania kół osi przedniej w celu zapewnienia stateczności kursowej:

$$f_1 > z > f_2 \quad \text{dla} \quad z = 0{,}15 \div 0{,}30 \tag{3}$$

Drugie rozwiązanie: wskaźniki przyczepności wykorzystanej przez obie osie powinny mieścić się w określonym paśmie, a wówczas granice blokowania kół określone są zależnościami:

$$\begin{matrix} f_1 \ge z - 0{,}08 \\ f_{1,2} \le z + 0{,}08 \end{matrix} \quad \text{dla} \quad z = 0{,}15 \div 0{,}30 \tag{4}$$

Ponadto krzywa wykorzystania przyczepności dla osi tylnej powinna spełniać warunek:

$$f_2 \le \frac{z-0.02}{0.74} \quad \text{for} \quad z \ge 0.3 \tag{5}$$

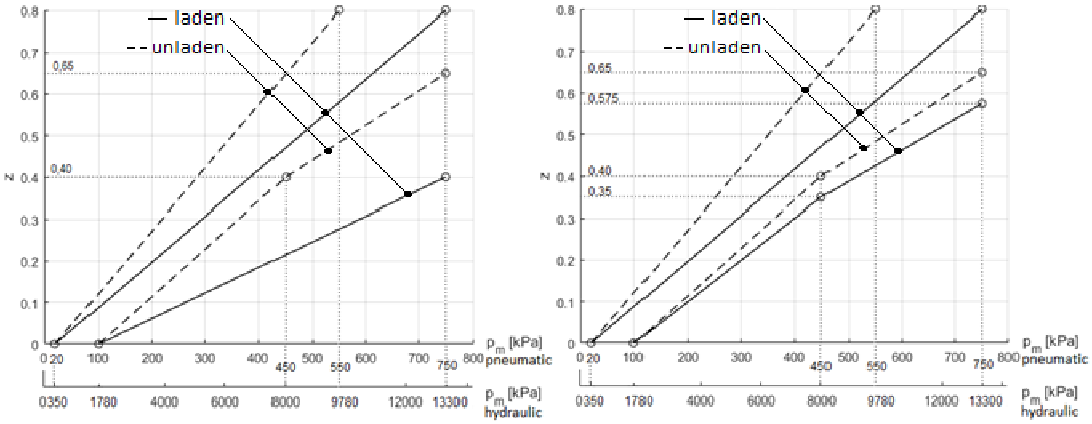
Opisane powyżej wymagania odnoszą się również do przyczep z większą niż dwie liczbą osi. Wówczas wskaźniki przyczepności wykorzystanej przez zespół osi przednich i zespół osi tylnych oblicza się z zależności:

$$f_1 = \frac{\sum f_{1i} R_{1i}}{\sum R_{1i}} \qquad f_2 = \frac{\sum f_{2i} R_{2i}}{\sum R_{2i}} \tag{6}$$

Wymogi dotyczące kolejności blokowania kół uznaje się za spełnione jeżeli dla wskaźników skuteczności hamowania w zakresie od 0,15 do 0,30 przyczepność wykorzystana przez co najmniej jedną z przednich osi jest większa niż przyczepność wykorzystana przez co najmniej jedną z osi tylnych [2]:

$$f_{1i} > f_{2i} \quad \text{dla dowolnego } i \tag{7}$$

W rozważaniach dotyczących wytycznych rozdziału sił hamujących (2)-(5) każda część zestawu drogowego traktowana jest jako pojedynczy pojazd, bez uwzględnienia sterowania hamulcami pojazdów ciągniętych. Stąd też w celu zapewnienia kompatybilności sił hamowania w zespole pojazdów wyznaczono dopuszczalne pasma zmian wskaźników hamowania poszczególnych pojazdów dla ich krańcowych stanów obciążenia w funkcji ciśnienia sterującego na głowicy sprzęgającej [2] – rys 2.



Rys.2. Dopuszczalne pasma wskaźnika hamowania dla traktorów z_M i przyczep z_R w funkcji ciśnienia p_m w przewodzie sterującym

3. Wyznaczanie dopuszczalnego obszaru rozdziału sił hamujących w przyczepach dwuosiowych

Jak wynika z rys. 3 naciski kół osi przedniej i kół osi tylnej przyczepy na nawierzchnię drogi poziomej zmieniają się zależnie od intensywności hamowania (wskaźnika skuteczności hamowania z) następująco:

$$R_1 = \frac{G}{L}(b + h \cdot z) \qquad R_2 = \frac{G}{L}(L - b - h \cdot z) \tag{8}$$

gdzie: L – rozstaw osi przyczepy, h – wysokość położenia środka ciężkości nad podłożem, b – odległość środka ciężkości od płaszczyzny pionowej przechodzącej przez oś tylną.

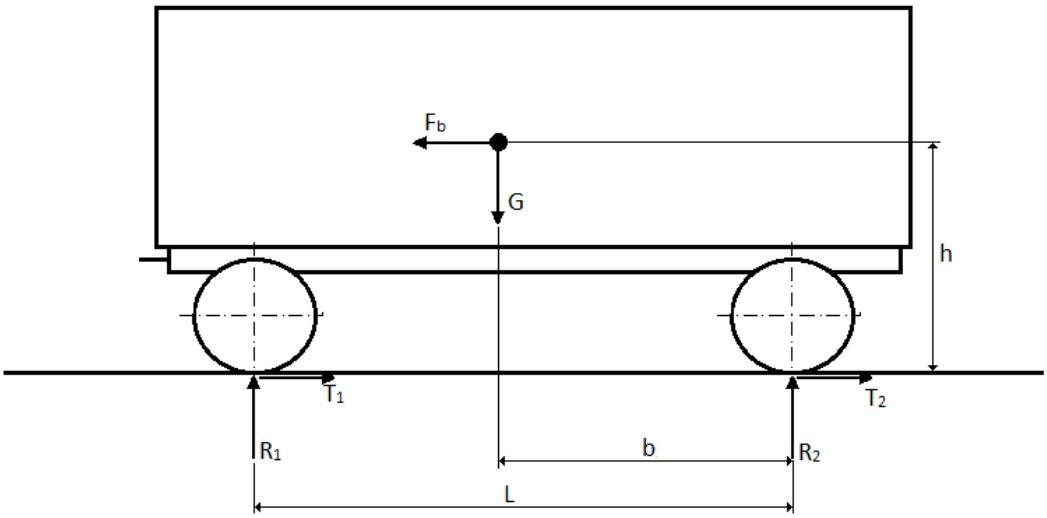


Fig. 3. Diagram showing forces acting upon a two-axle trailer during braking.

Względne (odniesione do ciężaru G przyczepy) siły hamowania osi przedniej γ_1 i tylnej γ_2 oblicza się z zależności:

$$\gamma_1 = \frac{T_1}{G} = \frac{R_1 f_1}{G} = \left(\frac{b}{L} + \frac{h}{L} z\right) f_1 \qquad \gamma_2 = \frac{T_2}{G} = \frac{R_2 f_2}{G} = \left(1 - \frac{b}{L} - \frac{h}{L} z\right) f_2 \tag{9}$$

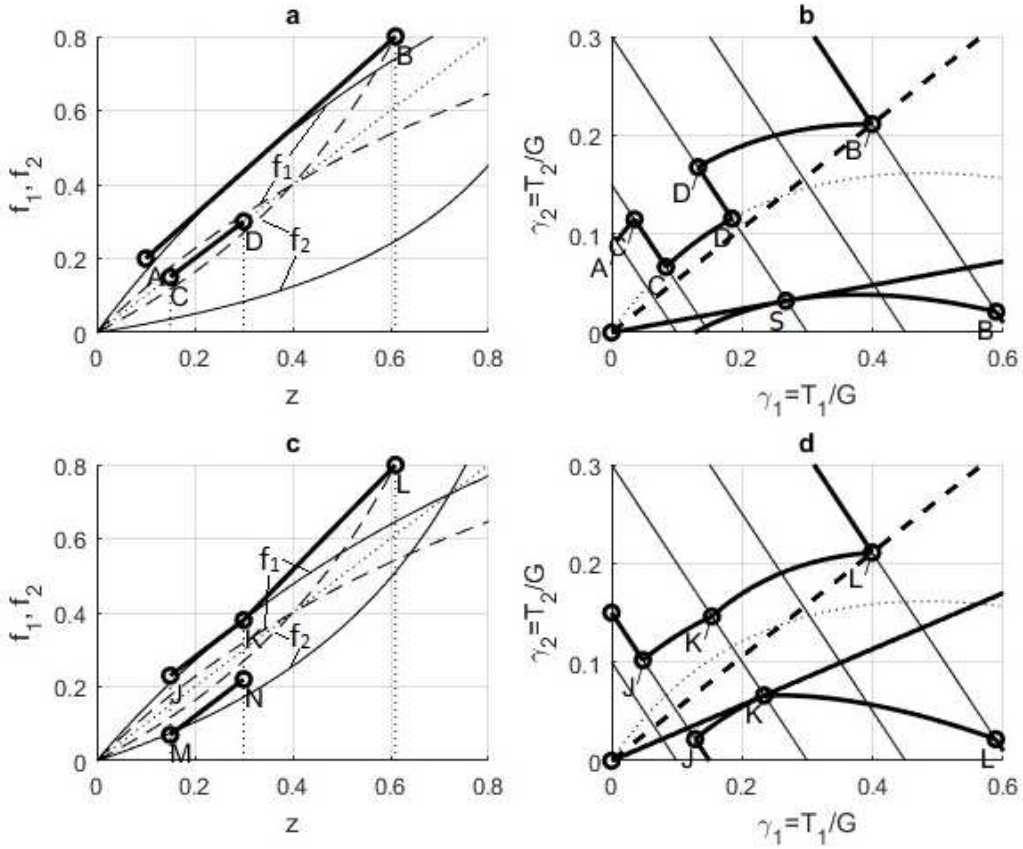
gdzie: T_1, T_2 – siły hamowania osi przedniej i tylnej, R_1, R_2 – reakcje pionowe nawierzchni na oś przednią i tylną pojazdu:

W warunkach hamowania idealnego wskaźniki przyczepności wykorzystanej przez przednią i tylną oś przyczepy są jednakowe i równe intensywności hamowania $f_1=f_2=z$, a rozdział sił hamujących jest opisany równaniem parametrycznym:

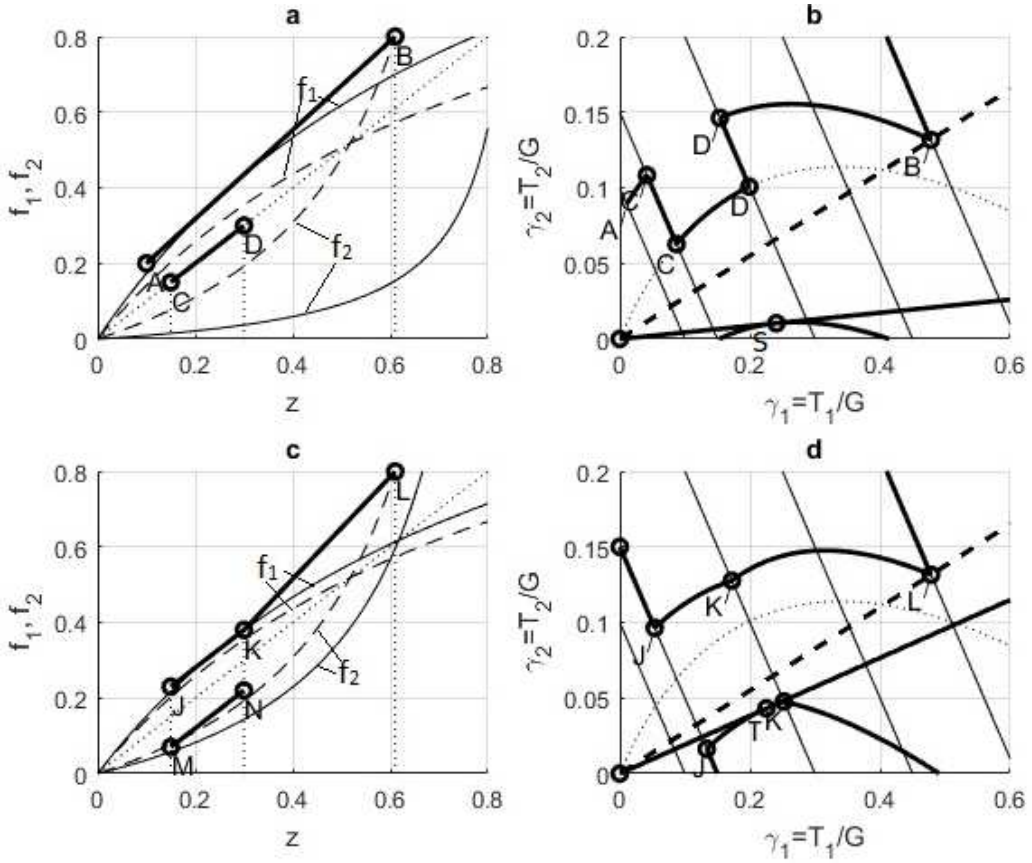
$$\gamma_1 = \left(\frac{b}{L} + \frac{h}{L} z\right) z \qquad \gamma_2 = \left(1 - \frac{b}{L} - \frac{h}{L} z\right) z \tag{10}$$

Wykorzystując zależności na graniczne wartości współczynników przyczepności wykorzystanej przez osie (2,3,4,5) oraz dane techniczne przyczepy można wyznaczyć dolną i górną granicę dopuszczalnego rozdziału sił hamujących na wykresie względnych sił hamujących γ_2=f(γ_1). Graficzną interpretację opisanych zaleceń wg pierwszego rozwiązania obrazują linie AB i CD na rys.4-a i rys.5-a. Odpowiadające im ograniczenia sił hamujących w układzie współrzędnych γ_1-γ_2 dla przykładowej przyczepy w stanie pustym i załadowanym przedstawiono na rys.4-b i rys.5-b. Krzywe graniczne oblicza się podstawiając wyznaczone z warunków (2), (3) wskaźniki przyczepności f_1, f_2 do zależności (9).

W drugim rozwiązaniu granice obszaru dopuszczalnego wskaźników przyczepności wytyczają linie MN i JKL na rys.4-c i rys.5-c. Odpowiadające im obszary względnych sił hamowania wg drugiego rozwiązania pokazano na rys.4-d dla przyczepy pustej, a dla przyczepy załadowanej na rys.5-d. Ze względu na restrykcyjny charakter warunku (4) dla górnej granicy K'L' na wykresie γ_2=f(γ_1) ograniczono zakres jego obowiązywania do przedziału z=0,3÷0,61.



Rys.4. Wyznaczanie parametrów stałego rozdziału sił hamujących dla przyczepy pustej o masie 4200 kg: a, c – przebiegi wskaźników przyczepności wykorzystanej przez osie f_1, f_2; b – wartości graniczne współczynnika rozdziału wg rozwiązania 1; d – wartości graniczne współczynnika rozdziału wg rozwiązania 2; L=2,95 m; b=1,47m; h=1,15 m



Rys.5. Wyznaczanie parametrów stałego rozdziału sił hamujących dla przyczepy załadowanej o masie 16250 kg: a, c – przebiegi wskaźników przyczepności wykorzystanej przez osie f_1, f_2; b – wartości graniczne współczynnika rozdziału wg rozwiązania 1; d – wartości graniczne współczynnika rozdziału wg rozwiązania 2; L=2,95 m; b=1,47m; h=1,63 m

Równania poszczególnych linii i krzywych granicznych w układzie $f_{1,2}$-z i γ_1-γ_2 wraz ze współrzędnymi poszczególnych punktów zestawiono w tab.2 i tab.3.

Tabela. 2. Wymagania skuteczności i stateczności hamowania dla przyczep

Krzywa	W układzie $f_{1,2}$-z	W układzie γ_1-γ_2	Zakres
A-B	$z \geq 0{,}85 \cdot f_{1,2} - 0{,}07$	$\gamma_1 = \left(\frac{b}{L} + \frac{h}{L} z\right)\left(\frac{z + 0{,}07}{0{,}85}\right)$	z=0,1-0,61
C-D	$z \leq f_{1,2}$	$\gamma_1 = \left(\frac{b}{L} + \frac{h}{L} z\right) z$	z=0,15-0,30
A'-C' D'-B'	$z \geq 0{,}85 \cdot f_{1,2} - 0{,}07$	$\gamma_1 = z - \left(1 - \frac{b}{L} - \frac{h}{L} z\right)\left(\frac{z + 0{,}07}{0{,}85}\right)$	z=0,1-0,15 z=0,3-0,61
J-K	$z \leq f_{1,2} - 0{,}08$	$\gamma_1 = \min\begin{cases} \left(\frac{b}{L} + \frac{h}{L} z\right)(z + 0{,}08) \\ z - \left(1 - \frac{b}{L} - \frac{h}{L} z\right)(z - 0{,}08) \end{cases}$	z=0,15-0,30

M-N	$z \le f_{1,2} + 0{,}08$	$\gamma_1 = \max\begin{cases} \left(\frac{b}{L}+\frac{h}{L}z\right)(z-0{,}08) \\ z-\left(1-\frac{b}{L}-\frac{h}{L}z\right)(z+0{,}08) \end{cases}$	z=0,15-0,30
K-L	$z \ge 0{,}3+0{,}74(f_{1,2}-0{,}38)$	$\gamma_1 = \left(\frac{b}{L}+\frac{h}{L}z\right)\left(\frac{z-0{,}3}{0{,}74}+0{,}38\right)$	z=0,30-0,61
K'-L'	$z \ge 0{,}3+0{,}74(f_{1,2}-0{,}38)$	$\gamma_1 = z-\left(1-\frac{b}{L}-\frac{h}{L}z\right)\left(\frac{z-0{,}3}{0{,}74}+0{,}38\right)$	z=0,30-0,61

Tabela. 3. Współrzędne punktów charakterystycznych; dla wszystkich zakresów obowiązuje zależność $\gamma_2 = z - \gamma_1$

Punkt	z	$f_{1,2}$	γ_1
A	0,10	0,20	$0{,}2(b/L+0.1\cdot h/L)$
B	0,61	0,80	$0{,}8(b/L+0.61\cdot h/L)$
C	0,15	0,15	$0{,}15(b/L+0.15\cdot h/L)$
D	0,30	0,30	$0{,}3(b/L+0{,}3\cdot h/L)$
A'	0,10	0,20	$0{,}1-0{,}2(1-b/L-0{,}1\cdot h/L)$
C'	0,15	0,259	$0{,}15-(0{,}22/0{,}85)(1-b/L-0.15\cdot h/L)$
D'	0,30	0,435	$0{,}3-(0{,}37/0{,}85)(1-b/L-0{,}3\cdot h/L)$
B'	0,61	0,8	$0{,}61-(0{,}68/0{,}85)(1-b/L-0{,}61\cdot h/L)$
J	0,15	0,23	$\min\begin{cases} 0{,}23(b/L+0{,}15\cdot h/L) \\ 0{,}15-0{,}07(1-b/L-0{,}15\cdot h/L) \end{cases}$
K	0,30	0,38	$\min\begin{cases} 0{,}38(b/L+0{,}3\cdot h/L) \\ 0{,}3-0{,}22(1-b/L-0{,}3\cdot h/L) \end{cases}$
L	0,61	0,8	$(b/L+0{,}61\cdot h/L)(0{,}31/0{,}74+0{,}38)$
M	0,15	0,07	$\max\begin{cases} 0{,}07(b/L+0{,}15\cdot h/L) \\ 0{,}15-0{,}23(1-b/L-0{,}15\cdot h/L) \end{cases}$
N	0,30	0,22	$\min\begin{cases} 0{,}22(b/L+0{,}3\cdot h/L) \\ 0{,}3-0{,}38(1-b/L-0{,}3\cdot h/L) \end{cases}$
K'	0,30	0,38	$0{,}38(1-b/L-0{,}3\cdot h/L)$
L'	0,61	0,8	0,61- $(1-b/L-0{,}61\cdot h/L)(0{,}31/0{,}74+0{,}38)$

4. Dobór liniowego rozdziału sił hamujących w przyczepach dwuosiowych

W powietrznych układach hamulcowych przyczep rolniczych zazwyczaj stosowane są korektory rozdziału sił hamujących o charakterystyce promienistej (liniowej) [8, 17, 28]. Charakterystyka ta jest opisana równaniem linii prostej przechodzącej przez początek układu współrzędnych i drugi wybrany punkt na wykresie względnych sił hamujących γ_2=f(γ_1) z uwzględnieniem ograniczeń obszarowych opisanych w poprzednim rozdziale. Procedurę wyznaczania dopuszczalnego zakresu zmian współczynnika

kierunkowego $i_p=T_2/T_1=\gamma_2/\gamma_1$ prostych obrazujących stały rozdział sił hamujących należy przeprowadzić dla pojazdu pustego i załadowanego.

W pierwszym rozwiązaniu obszar dopuszczalnego liniowego rozdziału sił hamujących wyznacza od dołu prosta OS styczna do krzywej granicznej AB w punkcie S (rys.4-b, rys.5-b), a od góry prosta przechodząca przez punkt D lub B' (wybiera się prostą o mniejszej wartości współczynnika kierunkowego).

Gdy korzystamy z drugiego rozwiązania, to dolną granicę obszaru dopuszczalnego wyznacza linia prosta styczna w punkcie T z krzywą JK (rys.4-d). Jeżeli punkt styczności T leży poza wycinkiem JK krzywej granicznej, to współczynnik kierunkowy prostej granicznej wyznacza się na podstawie współrzędnych punktu K (rys.5-d). Górną granicę liniowego rozdziału sił hamujących wyznacza prosta przechodząca przez punkt L' (rys.4-d, rys.5-d). Niekiedy może to być również punkt N.

Ponieważ prosta liniowego rozdziału sił hamowania przechodzi przez początek układu współrzędnych $\gamma_2=f(\gamma_1)$, to jej współczynnik kierunkowy jest w każdym przypadku obliczany ze stosunku rzędnej do odciętej danego punktu charakterystycznego *P*:

$$i_P=\frac{\gamma_{2p}}{\gamma_{1p}}=\frac{1-b/L-z_p\cdot h/L}{b/L+z_p\cdot h/L} \tag{11}$$

Gdzie: *P* – symbol punktu charakterystycznego.

Wykorzystując zależność $z=\gamma_1+\gamma_2$, można dla danej linii opisać rozdział sił hamujących poszczególnych osi za pomocą równania parametrycznego:

$$T_1=\frac{1}{1+i_P}G\cdot z \qquad T_2=\frac{i_P}{1+i_P}G\cdot z \tag{12}$$

w którym parametrem jest wskaźnik hamowania *z*. Wskaźniki przyczepności wykorzystanej przez osie na danej linii rozdziału sił hamujących oblicza się ze wzoru:

$$f_1=\frac{T_1}{R_1}=\frac{z}{\left(\frac{b}{L}+z\frac{h}{L}\right)(1+i_P)} \qquad f_2=\frac{T_2}{R_2}=\frac{i_p\cdot z}{\left(1-\frac{b}{L}-z\frac{h}{L}\right)(1+i_p)} \tag{13}$$

Wyniki obliczeń granicznych wartości współczynników kierunkowych i_P dla rozpatrywanej przyczepy dwuosiowej zestawiono w tabeli 4.

Tabela 4. Wartości graniczne współczynników kierunkowych liniowego rozdziału sił hamujących dla przyczepy dwuosiowej

Wariant rozwiązania	Przyczepa pusta	Przyczepa załadowana
Pierwszy wariant wg (2), (3)	$i_{min}=i_S=0{,}1202$	$i_{min}=i_S=0{,}0434$
	$i_{max}=i_{B'}=0{,}5293$	$i_{max}=i_{B'}=0{,}2754$
Drugi wariant wg (4), (5)	$i_{min}=i_K=0{,}2832$	$i_{min}=i_T=0{,}1914$
	$i_{max}=i_{L'}=0{,}5383$	$i_{max}=i_{L'}=0{,}2750$
Współczynnik optymalny wg (16)	$i_{opt}=0{,}5759$	$i_{opt}=0{,}4463$

Obliczone wg zależności (13) przebiegi wskaźników f_1, f_2 przyczepności wykorzystanej przez osie, odpowiadające poszczególnym liniom granicznym rozdziału sił hamujących (tab.4) dla 1 i 2 rozwiązania naniesiono na rys.4,5-a,c.

Dobierając zakres zmian rzeczywistego rozdział sił hamujących $i=T_2/T_1$ dla przyczepy pustej i załadowanej należy dążyć do zbliżenia z górną prostą graniczną. Zapewnia to krótką drogę hamowania przy jednoczesnym niebezpieczeństwie wcześniejszego blokowania kół tylnych w zakresie intensywności większych od wartości wskaźnika hamowania w punkcie przecięcia krzywej rozdziału idealnego (10) z linią rozdziału stałego.

Miarą stopnia efektywności dobranego rozdziału sił hamujących pojazdu dla różnych wartości μ jest współczynnik wykorzystania przyczepności (współczynnik wykorzystania ciężaru pojazdu na hamowanie):

$$\zeta(\mu) = \frac{T}{\mu \cdot G} = \frac{z}{\mu} \tag{14}$$

Gdzie: μ - współczynnik przyczepności przylgowej.

W poszukiwaniu optymalnej wartości współczynnika kierunkowego prostej liniowego rozdziału sił hamujących wykorzystuje się kryterium równości minimalnych współczynników wykorzystania przyczepności dla dwóch skrajnych wartości współczynników przyczepności $\mu_1 < \mu < \mu_2$ [6] charakteryzujących warunki eksploatacji pojazdu:

$$\zeta(\mu_1) = \zeta(\mu_2) \tag{15}$$

lub kryterium maksymalizacji średniej wartości współczynnika wykorzystania przyczepności $\zeta(\mu)$ w zadanym przedziale (μ_1, μ_2) [7]:

$$\zeta_{sr} = \frac{1}{\mu_2 - \mu_1} \int_{\mu_1}^{\mu_2} \zeta(\mu) d\mu \tag{16}$$

W przypadku przyczep dwuosiowych dla obu kryteriów otrzymuje się jednakową optymalną wartość współczynnika przyczepności [10]:

$$\mu_{op} = \mu_1 + \frac{b}{L}(\mu_2 - \mu_1) \tag{17}$$

Na tej podstawie łatwo można określić optymalną wartość współczynnika kierunkowego rozdziału sił hamujących [9,10]:

$$i_{opt} = \frac{1 - b/L - \mu_{op} \cdot h/L}{b/L + \mu_{op} \cdot h/L} \tag{18}$$

zmieniając odpowiednio wartości b/L i h/L dla przyczepy pustej i załadowanej. W obliczeniach dla przyczep rolniczych można przyjmować μ_1=0,2 i μ_2=0,5 [9]. Wyliczona z zależności (18) optymalna linia stałego rozdziału sił hamujących musi leżeć w obszarze dopuszczalnym, wytyczonym przez proste graniczne (tab.4). W rozpatrywanym przypadku optymalne wartości współczynnika kierunkowego dla przyczepy pustej i załadowanej są większe od maksymalnej dopuszczalnej wartości. Niemniej, fakt ten przemawia za przyjmowaniem większych wartości współczynników rozdziału sił hamowania, zbliżonych do wartości optymalnych (drugie rozwiązanie).

W przypadku, gdy linie rozdziału względnych sił hamujących przechodzą przez punkt B' (pierwsze rozwiązanie) lub L' (drugie rozwiązanie), to współczynniki rozdziału są identyczne. W obu przypadkach krzywa $f_2(z)$ przyczepności wykorzystanej przez oś tylną przechodzi przez punkt o współrzędnych z=0,61 i f_2(0,61)=0,8. Wyliczając siły hamowania dla tego punktu:

$$\begin{aligned} T_2 &= f_2 R_2 = 0{,}8G\left(1 - \frac{b}{L} - 0{,}61\frac{h}{L}\right) \\ T_1 &= G \cdot z - T_2 = G\left[0{,}61 - 0{,}8\left(1 - \frac{b}{L} - 0{,}61\frac{h}{L}\right)\right] \end{aligned} \tag{19}$$

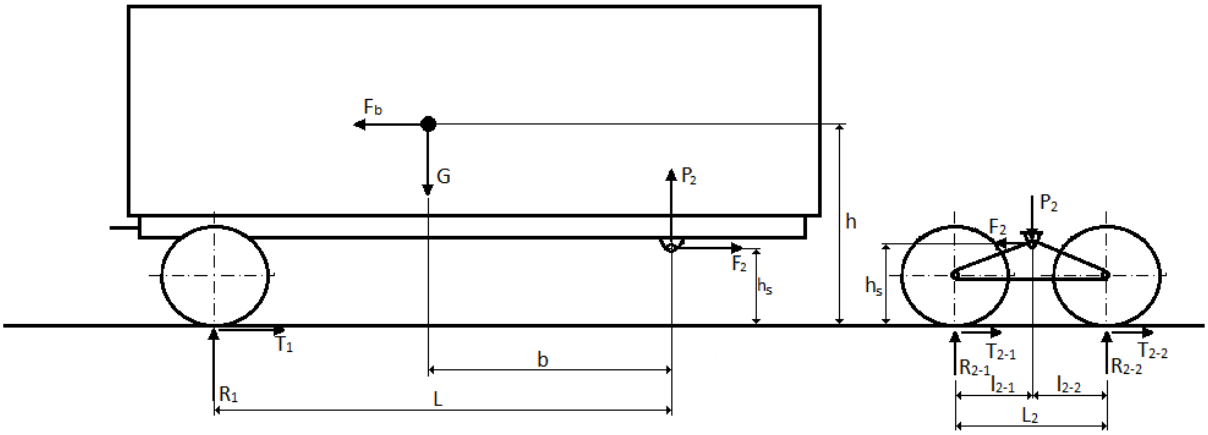
otrzymuje się następujące wyrażenie na współczynnik rozdziału sił hamujących:

$$i_P = \frac{T_2}{T_1} = \frac{0{,}8(1 - b/L - 0{,}61h/L)}{0{,}61 - 0{,}8(1 - b/L - 0{,}61h/L)} \tag{20}$$

Różnice między maksymalnymi wartościami i_{max} podanymi w tab. 4 dla obu rozwiązań wynikają z zaokrągleń współczynników nierówności (5). Dla dokładnych obliczeń dzielnik w wyrażeniu (5) powinien wynosić 0,7381.

5. Dobór liniowego rozdziału sił hamujących w przyczepach trzyosiowych

W rolniczych przyczepach trzyosiowych dwie osie tylne są umiejscowione blisko i pracują w układzie tandem. Przypadającą na zespól osi tylnych siłę hamowania należy rozdzielić stosownie do rozkładu obciążenia między osiami zawieszenia tandemowego. Układ sił działających na rolniczą przyczepę trzyosiową z zawieszeniem tandem typu belkowego „walking beam" przedstawiono na rys.6. W przyjętym modelu obliczeniowym zakłada się pominięcie masy nieresorowanej osi tandem, co oznacza pominięcie sił ciężkości i sił bezwładności zawieszenia.



Rys.6. Schemat sił działających na przyczepę trzyosiową z zawieszeniem tandem typu „walking beam"

Bilans sił i momentów działających na przyczepę ma postać:

$$\sum F_x = T_1 + T_2 - G \cdot z = 0 \tag{21}$$

$$\sum F_y = R_1 + R_2 - G = 0 \tag{22}$$

$$\sum M_2 = G \cdot b + G \cdot z(h - h_s) + T_1 \cdot h_s - R_1 \cdot L = 0 \tag{23}$$

Po rozwiązaniu powyższego układu równań otrzymuje się zależności opisujące reakcje R_1 i R_2 działające na przyczepę:

$$R_1 = G\left(\frac{b}{L} + z\frac{h - h_s}{L}\right) + T_1\frac{h_s}{L} \quad \text{lub} \quad R_1 = G\left(\frac{b}{L} + z\frac{h}{L}\right) - T_2\frac{h_s}{L} \tag{24}$$

$$R_2 = G\left(1 - \frac{b}{L} - z\frac{h - h_s}{L}\right) - T_1\frac{h_s}{L} \quad \text{lub} \quad R_2 = G\left(1 - \frac{b}{L} - z\frac{h}{L}\right) + T_2\frac{h_s}{L} \tag{25}$$

Gdzie:

$$T_1 = \frac{1}{1 + i_P} G \cdot z \qquad T_2 = G \cdot z - T_1 \tag{26}$$

Bilans sił i momentów działających na zawieszenie tandemowe ma postać:

$$\sum F_x = T_{21} + T_{22} - T_2 = 0 \tag{27}$$

$$\sum F_y = R_{21} + R_{22} - R_2 = 0 \tag{28}$$

$$\sum M_{21} = T_2 \cdot h_s + R_{22} L_2 - R_2 \cdot L_{21} = 0 \tag{29}$$

W celu wyznaczenia rozdziału siły hamowania T_2 pomiędzy koła zespołu osi tandem przyjęto liniowy rozdział sił hamujących:

$$\frac{T_{22}}{T_{21}} = i_S \tag{30}$$

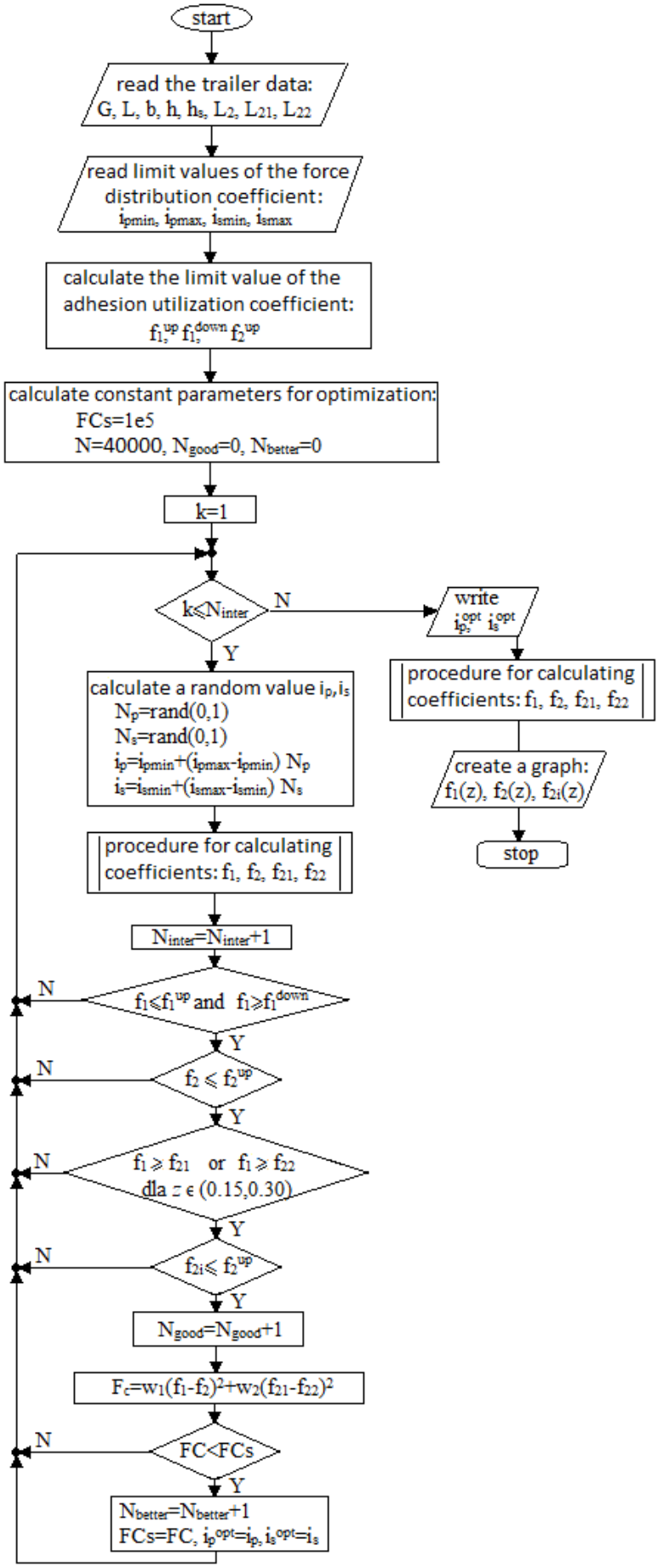
Rozwiązując układ równań (27)-(29) otrzymujemy siły działające na koła zawieszenia tandemowego podczas hamowania:

$$R_{21} = R_2 \frac{L_{22}}{L_2} + T_2 \frac{h_s}{L_2} \qquad R_{22} = R_2 \frac{L_{21}}{L_2} - T_2 \frac{h_s}{L_2} \tag{31}$$

$$T_{21} = \frac{T_2}{1+i_S} \qquad T_{22} = \frac{i_S}{1+i_S} T_2 \tag{32}$$

Zależności (24), (25) oraz (31), (32) można stosować również dla zawieszenia tandemowego typu „boogie”.

Do przeszukiwania dopuszczalnego obszaru zmienności współczynników i_P i i_S rozdziału sił hamujących w celu znalezienia rozwiązań optymalnych zastosowano metodę Monte Carlo [3], [18], [20]. Schemat blokowy algorytmu optymalnego doboru współczynników rozdziału sił hamujących pokazano na rys.7. Na jego podstawie opracowano program komputerowy w środowisku Matlab [25].


start
read the trailer data:
G, L, b, h, hs, L2, L21, L22
read limit values of the force distribution coefficient:
ipmin, ipmax, ismin, ismax
calculate the limit value of the adhesion utilization coefficient:
f1,up f1,down f2up
calculate constant parameters for optimization:
FCs=1e5
N=40000, Ngood=0, Nbetter=0
k=1
k≤Ninter
N
Y
write
ipopt isopt
procedure for calculating coefficients: f1, f2, f21, f22
create a graph:
f1(z), f2(z), f2i(z)
stop
calculate a random value ip,is
Np=rand(0,1)
Ns=rand(0,1)
ip=ipmin+(ipmax-ipmin) Np
is=ismin+(ismax-ismin) Ns
procedure for calculating coefficients: f1, f2, f21, f22
Ninter=Ninter+1
f1≤f1up and f1≥f1down
f2 ≤ f2up
f1 ≥ f21 or f1 ≥ f22
dla z ϵ (0.15,0.30)
f2i≤ f2up
Ngood=Ngood+1
Fc=w1(f1-f2)2+w2(f21-f22)2
FC<FCs
Nbetter=Nbetter+1
FCs=FC, ipopt=ip, isopt=is

Rys.7. Schemat blokowy algorytmu optymalizacji rozdziału sił hamujących przyczepy trzyosiowej metodą Monte Carlo (FC_s – wartość początkowa funkcji celu, N – liczba losowań, N_{good} – liczba rozwiązań dobrych, spełniających ograniczenia nierównościowe, N_{better} – liczba rozwiązań lepszych, zmniejszających wartość funkcji celu)

Optymalne wartości współczynników rozdziału sił hamujących ustalano w procesie minimalizacji funkcji celu w postaci:

$$FC = \frac{w_1(f_1 - f_2)^2 + w_2(f_{21} - f_{22})^2}{w_1 + w_2} \tag{33}$$

gdzie: w_i – współczynniki wagowe.
Tak sformułowana funkcja preferuje rozwiązania zbliżające do siebie współczynniki f_i przyczepności wykorzystanej przez poszczególne osie.

Przed obliczeniem funkcji celu sprawdzane są ograniczenia nierównościowe (4), (5) dla rozwiązania drugiego:

$$\begin{aligned} &f_1 \geq f_1^{down} = z - 0{,}08 \\ &f_1 \leq f_1^{up} = z + 0{,}08 \end{aligned} \quad \text{dla} \quad z = 0{,}15 \div 0{,}30$$
$$f_2 \leq f_2^{up} = (z + 0{,}08)(0.15 \leq z \leq 0{,}30) + \left(\frac{z - 0{,}3}{0{,}7381} + 0{,}38\right)(z \geq 0.30) \tag{34}$$

oraz warunek (7):

$$f_1 > f_{21} \text{ or } f_1 > f_{22} \text{ dla } z = 0{,}15 \div 0{,}30 \tag{35}$$

Ponadto przyjęto dodatkowy warunek dla współczynników wykorzystania przyczepności osi tylnych:

$$f_{2i} \leq f_2^{up} \tag{36}$$

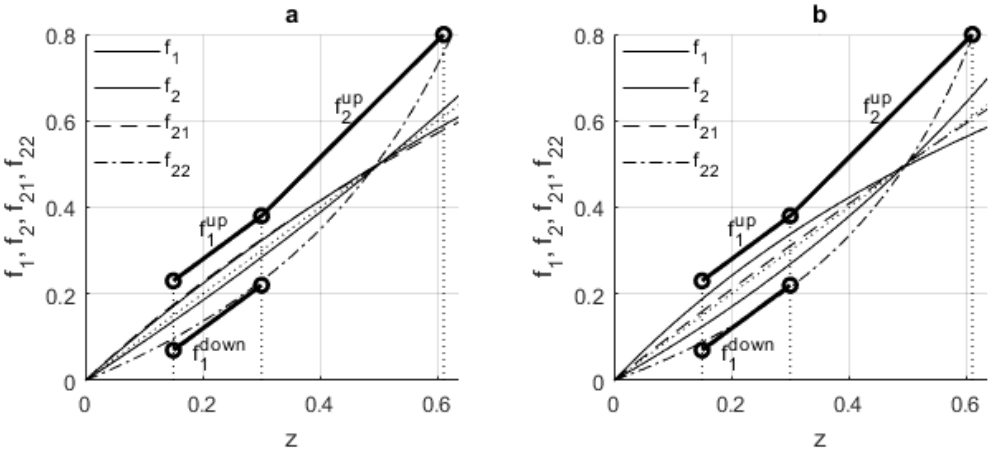
ograniczający nadmierny wzrost współczynnika f_{22} dla $z \leq 0{,}61$.

Wyniki obliczeń rozdziału sił hamowania dla przyczepy pustej i załadowanej po wykonaniu kilku uruchomień programu zestawiono w tab.5. Przyjęto liczbę losowań N=40000, w_1=0,6, w_2=0,4.

Tab.5. Wyniki optymalizacji rozdziału sił hamujących w przyczepie trzyosiowej

L.p.	1	2	3	4	5	Średnia
	Przyczepa pusta					
i_S	1.2971	1.2963	1.2944	1.2991	1.2955	1.2965
i_P	0.5150	0.5162	0.5155	0.5153	0.5157	0.5155
FC	1.7961	0.8805	0.4060	0.7386	1.5802	1.0803
	Przyczepa załadowana					
i_S	0.9753	0.9737	0.9749	0.9729	0.9725	0.9739
i_P	0.5326	0.5337	0.5314	0.5333	0.5331	0.5328
FC	0.5287	0.5616	0.3032	1.1820	0.6626	0.6476

Przykładowy przebiegi wskaźników wykorzystania przyczepności $f_i(z)$ przez osie dla optymalnego rozdziału siła hamujących dla przyczepy pustej i załadowanej pokazano na rys.8.



Rys. 8. Przebiegi $f_i(z)$ dla optymalnego rozdziału sił hamujących: a- dla przyczepy pustej (i_s=1,2994, i_P=0,5155, m=7700kg, L=5,15m, b=1,85m, h=1,8m, L_2=1,345m, $L21$=0,72m $L22$=0,62m, h_s=0,545m), b – dla przyczepy załadowanej (i_s=0,9749, i_P=0,5333, m=24000kg, L=5,15m, b=1,85m, h=1,8m, L_2=1,36m, L_{21}=0,73m, L_{22}=0,63m, h_s=0,514m)

6. Obliczenia i dobór mechanizmu hamulcowego i mechanizmu uruchamiania

Znając wartości sił hamowania poszczególnych osi oblicza się parametry konstrukcyjne ich mechanizmów hamulcowych i mechanizmów uruchamiania. Można tu skorzystać z następującej zależności na siłę hamowania i- tej osi [2]:

$$T_i = k \cdot (C - C_0) \cdot \eta \cdot BF / r_d + f_r R_i \tag{37}$$

gdzie: k- liczba siłowników na oś, C- moment na wałku rozpieraka generowany przez aktuator, C_0 – minimalny moment wałka rozpieraka niezbędny do wytworzenia mierzalnego momentu hamowania, η – sprawność mechaniczna, r_d – promień dynamiczny koła, f_r – współczynniki oporu toczenia kół, R_i – obciążenie kół i-tej osi, BF – współczynnik „Brake factor" zdefiniowany następująco [1]:

$$BF = \frac{C^* \cdot r_e}{2r_b} \tag{38}$$

gdzie: C^* współczynnik efektywności (współczynnik wzmocnienia wewnętrzny) mechanizmu hamulcowego [15], r_e – promień czynny bębna hamulcowego, r_b – skuteczny promień krzywki rozpieraka.

Moment na wałku rozpieraka jest iloczynem siły Th_a rozwijanej przez siłownik hamulcowy działającej na dźwignię rozpieraka o długości l:

$$C = Th_a l \tag{39}$$

Wykorzystując dane eksperymentalne producentów siłowników można siłę użyteczną na tłoczysku wyrazić za pomocą zależności:

$$Th_a = A \cdot p - B \tag{40}$$

gdzie: A, B – współczynniki eksperymentalne, p – ciśnienie w komorze siłownika.

7. Podsumowanie i wnioski

Opisane w pracy obliczenia umożliwiają dobór optymalnego liniowego rozdziału sił hamujących w procesie projektowania powietrznych układów hamulcowych dwu i trzyosiowych przyczep rolniczych, w których stosowane są korektory sił hamujących o charakterystyce promienistej. W obliczeniach rozdziału sił hamujących uwzględniono wymogi Dyrektywy UE 2015/68 [2] w zakresie skuteczności i stateczności hamowania. Obliczenia rozdziału sił hamujących stanowią podstawę obliczeń projektowych i umożliwiają

w następnym etapie procesu projektowania dobór parametrów osi hamowanych (mechanizmu hamulcowego, siłownika i mechanizmu uruchamiania) oraz charakterystyk zaworów hamulcowych.

Obliczenia rozdziału sił hamujących wykonane dla przyczepy dwuosiowej o ładowności około 16 ton, przeprowadzone dla dwóch dopuszczalnych wariantów rozwiązań rozdziału sił hamujących, przemawiają za stosowaniem w obliczeniach projektowych drugiego rozwiązania (wg wymogów (4), (5)). Wyliczone dla tego rozwiązania współczynniki przyczepności wykorzystanej przez osie są bardziej zbliżone do prostej obrazującej idealny rozdział sił hamujących, w którym współczynniki przyczepności wykorzystanej przez każdą z osi są takie same i równe wskaźnikowi hamowania.

W przypadku przyczepy dwuosiowej zakres dopuszczalnych zmian współczynnika liniowego rozdziału sił hamujących oraz jego wartość optymalną dla różnych stanów załadowania wyznacza się analitycznie na podstawie wykresu jednostkowych sił hamowania γ_2 (γ_1). Natomiast w przypadku przyczep trzyosiowych, w których trzeba dodatkowo dokonać podziału sił hamujących między osie zespołu tandem, szybsze rezultaty otrzymuje się stosując metody optymalizacji.

Opracowany algorytm poszukuje optymalnego rozdziału siła hamujących metodą Monte Carlo dla przyczep z zawieszeniem osi tylnych typu „walking beam" lub „bogie", który można łatwo zaadoptować do doboru rozdziału sił hamujących w przyczepach z innymi typami osi tandemowych poprzez zmianę bloku procedury wyliczania współczynników przyczepności wykorzystanej przez osie (inne zależności na reakcje kół).

Na podstawie przedstawionej metodyki można opracować zasady rozdziału sił hamujących przyczepy przy zastosowaniu korektorów sił hamowania o innej charakterystyce niż promienista (liniowa).

Acknowledgement

Projekt finansowany z programu Ministra Nauki i Szkolnictwa Wyższego pod nazwą „Regionalna Inicjatywa Doskonałości" w latach 2019-2022 numer projektu 011 / RID / 2018/19 kwota finansowania 12 000 000 PLN

Ministerstwo Nauki i Szkolnictwa Wyższego

Literatura

1. Andrew J D. Braking of Road Vehicles. Oxford: Butterworth-Heinemann, 2014.
2. Commission Delegated Regulation (EU) 2015/68 supplementing Regulation (EU) No 167/2013 of the European Parliament and of the Council with regard to vehicle braking requirements for the approval of agricultural and forestry vehicles, October 2014.
3. Dimov I T, Sean McKee S. Monte Carlo Methods for Applied Scientists. World Scientific Press, 2004.
4. Forrer P. Brake systems in agricultural and forestry vehicles, http://www.paul-forrer.ch (accessed 07 May 2019).
5. Glišović J, Lukić J, Vanja Šušteršič V, Ćatić D. Development of tractors and trailers in accordance with the requirements of legal regulations. In: 9th International Quality Conference, Center for Quality, Faculty of Engineering, University of Kragujevac, June 2015, paper no. 3504: 193-201.

6. Gredeskul A B. O normativach effektivnosti tormoženija avtomobilej. Avtomobilnaja promyšlennost 1963; 6: 14-16.
7. Gredeskul A B, Fedosov V M, Skutnev V M. Opredelenie parametrov tormoznoj sistemy s regulatorom tormoznych sil. Avtomobilnaja promyšlennost 1975; 6: 24-26.
8. Haldex, Agricultural trailer product catalogue. Europe, Edition 1, 2015.
9. Kamiński Z. Distribution of braking forces in two-axle agricultural trailers. Teka Kom. Mot. Energ. Roln. 2005; 5: 80-86.
10. Kaminski Z, Miatluk M. Brake systems of road vehicles. Calculations. Bialystok: Wydawnictwo Politechniki Bialostockiej, 2005.
11. Kamiński Z. Simulation and experimental testing of the pneumatic brake systems of agricultural vehicles. Bialystok: Oficyna Wydawnicza Politechniki Białostockiej, 2012.
12. Kamiński Z, Kulikowski K. Determination of the functional and service characteristics of the pneumatic system of an agricultural tractor with mechanical brakes using simulation methods. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2015; 17(3): 355–364.
13. Kamiński Z. Mathematical modelling of the trailer brake control valve for simulation of the air brake system of farm tractors equipped with hydraulically actuated brakes. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16(4): 637–643.
14. Keyser DE, Hogan K. Hydraulic brake systems and components for off-highway vehicles and equipment. National Fluid Power Association Technical Paper Series 1992; I 92-1.4: 1-9.
15. Keyser DE. Full power hydraulic brake actuation, circuit design considerations for off-highway vehicles and equipment. In: 10th International Conference on Fluid Power - the Future for Hydraulics, Brugge, Belgium, 5-7 April 1993, edited by N. Way. Mechanical Engineering Publications, London.
16. Khaled M, Mahmoud R. Theoretical and experimental investigations one new adaptive duo servo drum brake with high and constant brake shoe factor, university Paderborn, 2005.
17. Knorr-Bremse, Agricultural and forestry vehicles. Brake equipment catalogue, Y206317 - (EN - Rev. 001), 2015.
18. Kroese DP, Taimre T, Botev ZI. Handbook of Monte Carlo Methods. New York, 2011.
19. Lin M, Zhang W. Dynamic simulation and experiment of a full power hydraulic braking system. Journal of University of Science and Technology Beijing 2007; 29(10): 70-75.
20. Morton DP, Popova E. Monte Carlo simulations for stochastic optimization: Encyclopedia of Optimization. In: Floudas CA, Pardalos PM (eds) Monte Carlo simulations for stochastic optimization. Kluwer Academic Publishers, 2001; 1529-1537.
21. Radlinski RW, Flick MA. Tractor and trailer brake system compatibility. SAE Transactions; paper no. 861942, 1986.
22. Safim. Trailer brake valve, http://www.italgidravlika.ru/pdf_files/Safim/safim_11.pdf (accessed 15 May 2018).
23. Tang G, Zhao H, Wu J, Zhang Y. Optimization of Braking Force Distribution for Three-Axle Truck. SAE Technical Paper 2013-01-0414, 2013.
24. UN Economic Commission for Europe, ECE Regulation No. 13. Uniform provisions concerning the approval of vehicles of categories M, N and O with regard to braking, Geneva, Switzerland, 2001.
25. Venkataraman P. Applied Optimization with MATLAB Programming Wiley-Interscience. New York, 2001.
26. Wabco, FPB – Full Hydraulic Power Brake, Version 2/09, 2013.
27. Wabco, Off-highway. Overview technologies and products, Edition 2, Version 3, December 2016.
28. Wabco, Air braking system. Agriculture and forestry vehicles, Edition 11, Version 1, October 2017.

dr inż. Anna Borucka
Military University of Technology
gen. Sylwestra Kaliskiego 2, 01-476, Warsaw, Poland
E-mail: anna.borucka@wat.edu.pl

prof. dr hab. inż. Andrzej Niewczas
Motor Transport Institute,
Jagiellońska 80, 03-301, Warsaw, Poland
E-mail: andrzej.niewczas@its.waw.pl

dr Kamila Hasilova
University of Defence,
Kounicova 65, 662 10 Brno, Czech Republic
E-mail: kamila.hasilova@unob.cz

Prognozowanie gotowości pojazdów specjalnych na podstawie modelu semi -Markowa
Forecasting the readiness of special vehicles using the semi-Markov model

Keywords: *vehicle exploitation system, special vehicles, readiness, semi-Markov model.*

Abstract: The vehicle exploitation system, consisting of statistically identical objects that perform intervention tasks, not subject to systematic changes can be modelled as a stationary stochastic process. Such a model allows to determine the probabilistic indicators of current and boundary readiness of the system. This article presents the use of the semi-Markov process, based on three operating states: operation, ready-to-be-used and repair to study a transport system consisting of special vehicles. On the example of a sample consisting of police patrol cars, experimental studies of the intensity of fleet utilization, time of failure-free operation of vehicles were carried out and it was demonstrated that the examined transport system is characterized by a satisfactory, stationary readiness coefficient. The developmental possibilities of the presented modelling method were emphasized.

Keywords: *system eksploatacji samochodów, pojazdy specjalne, gotowość, model semi-Markowa.*

Streszczenie: System eksploatacji samochodów, które realizują zadania interwencyjne, niepodlegający systematycznym zmianom może być modelowany jako stacjonarny proces stochastyczny. Taki model pozwala wyznaczyć probabilistyczne wskaźniki bieżącej i granicznej gotowości systemu.
W niniejszym artykule, do modelowania systemu eksploatacji pojazdów specjalnych, wykorzystano proces semi-Markowa, oparty na trzech stanach eksploatacyjnych: użytkowania, postoju użytkowego i naprawy. Na przykładzie próby radiowozów policyjnych przeprowadzono doświadczalne badania intensywności użytkowania floty, czasu bezawaryjnej pracy pojazdów a także wykazano, że badany system transportowy charakteryzuje się zadowalającym, stacjonarnym współczynnikiem gotowości. Podkreślono rozwojowe możliwości przedstawionej metody modelowania.

1. Wstęp

Proces eksploatacji pojazdów użytkowych można analizować zarówno w przedsiębiorstwach transportu samochodowego, które funkcjonują w warunkach rynkowych, jak również w służbach ratowniczych i innych, odpowiedzialnych za bezpieczeństwo państwa, takich jak straż pożarna, wojsko, policja, pogotowie ratunkowe. W pierwszej grupie najważniejszym kryterium oceny jakości eksploatacyjnej pojazdu jest efektywność, zazwyczaj wymiarowana

jako stosunek zysków do kosztów [1]. Druga grupa, w tym szczególnie policja, utożsamiana jest przede wszystkim z zapewnieniem spokoju obywatelom, ochroną życia i zdrowia ludzi oraz mienia, a także troską o porządek. Dlatego większość badań prowadzonych w tym obszarze tematycznym związana jest głównie z szeroko pojętymi kwestiami bezpieczeństwa i dotyczy na przykład:

1. Szacowania prawdopodobieństwa wypadku śmiertelnego w przypadku poruszania się samochodami policyjnymi [3] oraz oceny ryzyka zdarzeń drogowych w tym występowania poważnych urazów wynikających z uczestnictwa w akcjach policyjnych [6, 25].
2. Możliwości zwiększenia poziomu bezpieczeństwa realizacji akcji policyjnych poprzez zastosowanie specjalnych metod czy urządzeń np. proponowanych przez Michaelson'a kuloodpornych paneli montowanych na radiowozach [27] czy opisywanych przez Lyons'a systemów świateł ostrzegawczych [26].
3. Metod planowania i optymalizacji tras patroli [8, 10], ze szczególnym uwzględnieniem kwestii bezpieczeństwa [4], a także niezbędnej liczby samochodów patrolowych w zależności od natężenia realizowanych czynności i pory dnia ich występowania [22].

Natomiast gotowość i niezawodność pojazdów policyjnych uważana jest za swoiste status quo. Prezentowane w literaturze badania dotyczące oceny gotowości złożonych systemów interwencyjnych (nie tylko policji), mają charakter jednostkowy. Wynika to przede wszystkim z ograniczeń związanych z poufnym charakterem danych empirycznych. Dokumentacja ewidencyjno – rozliczeniowa prowadzona jest zazwyczaj w postaci papierowej a praktyka tworzenia elektronicznych baz danych, napotyka na bariery organizacyjne.

Zadania transportowe są procesami złożonymi, co powoduje że ich modelowanie w oparciu o klasyczne techniki teorii niezawodności może być zawiłe i nie dawać satysfakcjonujących wyników [21]. Wówczas wykorzystywane są alternatywne metody, np. proponowane przez Lu i współautorów [24] lub Dong'a i innych [9] niezawodnościowe diagramy fazowe, a także procesy Markowa [11, 16, 34], które w ocenie gotowości są szczególnie popularne. W literaturze można odnaleźć modele opisujące pojedyncze środki transportu, np. samochód osobowy – jak u Girtlera i Ślęzaka [12], autobus u Landowskiego i innych [23], czy śmigłowiec u Szawłowskiego [33]. Badane są także złożone systemy transportowe. Podstawy teoretyczne takich rozważań zawierają prace [2, 13, 20]. Systemy analizowane są jako całość [7, 32, 35] lub ich poszczególne składowe rozpatrywane są niezależnie, a każda z nich opisywana jest oddzielnym modelem. Często autorzy wskazują procesy Markowa jako narzędzie rozwiązania szczególnego problemu eksploatacyjnego [29, 30]. Niestety, modele systemów transportowych, oparte na danych empirycznych są nieliczne. Dostępne są pojedyncze opracowania, np. Migawa [28], badał w ten sposób system eksploatacji autobusów miejskich, Żurek i Tomaszewska [39] analizowali statki powietrzne, a Restel [31] systemy miejskiego transportu kolejowego.

Przegląd literatury pokazuje, że modele Markowa są dobrym narzędziem oceny gotowości zarówno całych systemów, jak i pojedynczych obiektów [5, 19]. Mają jednak swoje wymagania i ograniczenia. Należy do nich przede wszystkim postać dostępnych obserwacji, których rozkład powinien być wykładniczy. Jest to element często pomijany w prezentowanych analizach, co powoduje, że stosowanie procesów Markowa jest nadużywane. Trudniejsze w estymacji parametrów i dlatego mniej popularne są modele semi-Markowa. Mają one mniej restrykcyjne wymagania dotyczące postaci rozkładów badanych zmiennych (mogą być dowolne), dlatego zostały zaproponowane w niniejszym artykule, jako narzędzie oceny floty radiowozów policyjnych. Celem zaprezentowanego badania było oszacowanie poziomu ich gotowości przy założeniu trzech stanów eksploatacyjnych: użytkowania, postoju użytkowego i naprawy (obsługiwania technicznego) oraz przedstawienie metody stochastycznego opisu procesu eksploatacyjnego. Ponadto intencją

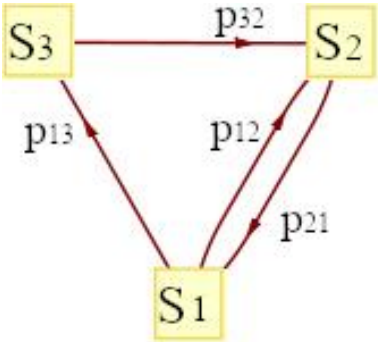
autorów było podkreślenie, że trzystanowy model eksploatacji może być użytecznym i wystarczającym narzędziem oceny gotowości pojazdów specjalnych. Zastosowanie takiego modelu nie wymaga skomplikowanych obliczeń jak ma to miejsce w przypadku rozbudowanych modeli kilkunastostanowych i może być wykorzystywane w bieżącej praktyce zarządzania flotą pojazdów.

2. Badania eksploatacyjne i analiza wstępna wyników

Przedmiotem badania były radiowozy policyjne, realizujące zadania patrolowe i interwencyjne na terenie miasta stołecznego Warszawy. Analizie poddano 20 oznakowanych samochodów osobowych marki Kia. Wszystkie samochody pochodziły z tej samej partii produkcyjnej, co pozwoliło uznać próbę za jednorodną. Bazę danych źródłowych stanowiła dokumentacja użytkowania radiowozów dotycząca patroli policyjnych oraz rejestry obsług technicznych i napraw.

Na podstawie zgromadzonych obserwacji wyodrębniono trzyelementowy zbiór stanów eksploatacyjnych $S = \{S_1, S_2, S_3\}$, w których przebywają pojazdy:
- użytkowanie (S_1),
- postój użytkowy (S_2),
- naprawa (również obsługiwania techniczne) (S_3).

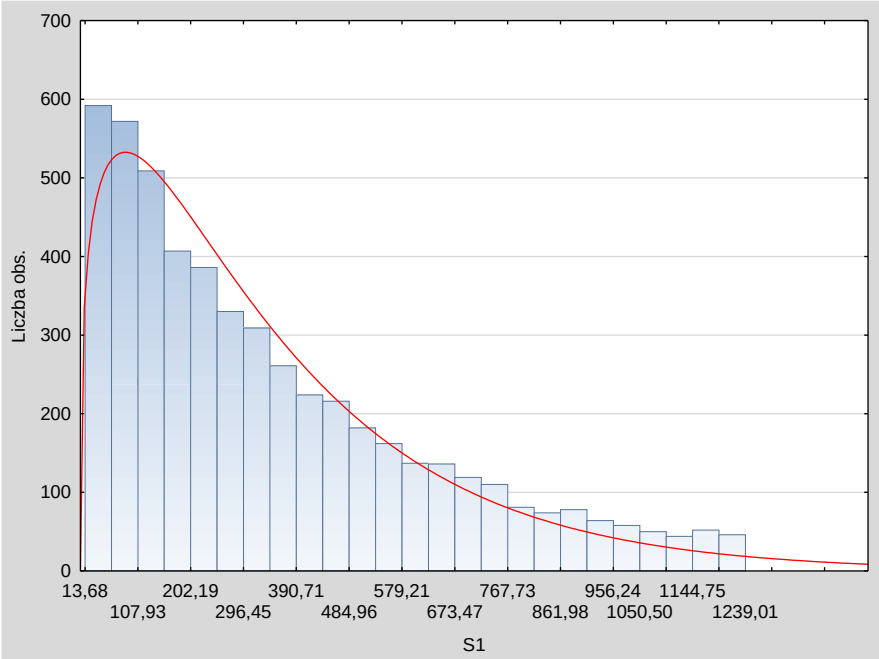
Przyjęto, że czas przebywania pojazdu w stanie S_1 (czas trwania stanu S_1) zawiera się w przedziale od chwili wyjazdu w celu wykonania zadania interwencyjnego (patrolu), do chwili powrotu pojazdu do zajezdni. Czas przebywania pojazdu w stanie S_2 (czas trwania stanu S_2) zawiera się od chwili rozpoczęcia postoju w zajezdni w oczekiwaniu na dyspozycję wykonania zadania do chwili wyjazdu. Czas przebywania pojazdu w stanie S_3 (czas trwania stanu S_3) wyznaczają chwile rozpoczęcia i zakończenia obsługiwania technicznego.
Następnie w oparciu o rzeczywiste relacje międzystanowe ustalono przejścia dozwolone, które, w formie grafu, zaprezentowano na rys. 1.



Rysunek 1. Graf przejść dozwolonych

Przeprowadzono także analizę statystycznych rozkładów czasów trwania (wyrażonych w minutach) poszczególnych stanów eksploatacyjnych. Sprawdzono dopasowanie rzeczywistych obserwacji do wybranych rozkładów teoretycznych (normalnego, log-normalnego, wykładniczego, Gamma i Weibulla). Parametry tych rozkładów estymowano z wykorzystaniem programu Statistica, wykorzystując metodę największej wiarogodności. Jakość dopasowania oceniono porównując rozkłady obserwowanych częstości z oczekiwanymi. Obliczono statystykę testu Kołmogorowa-Smirnowa oraz kryterium informacyjne Akaikego. Na podstawie otrzymanych wyników, jako najbardziej odpowiedni,

wybrano rozkład Gamma. Przykładową analizę przedstawiono dla rozkładu stanu użytkowanie - S_1 (rys. 2).



Rysunek 2 Histogram czasów trwania stanu S_1

3. Estymacja parametrów modelu semi-Markova

3.1 Podstawowe charakterystyki

Wnioskiem z przeprowadzonych analiz wstępnych było stwierdzenie braku możliwości zastosowania modelu Markowa (wymaga wykładniczej postaci rozkładów zmiennych) i założenie przeprowadzenia analiz z wykorzystaniem modelu semi -Markowa, dla którego postać rozkładów może być dowolna.

Dla badanego procesu eksploatacji samochodów określono model semi-Markowa o skończonym zbiorze stanów za pomocą markowskiego procesu odnowy, wzorując się na [12, 13, 20]: Dla N oznaczającego zbiór liczb całkowitych nieujemnych, S – pewien zbiór skończony, R_+ – zbiór liczb rzeczywistych nieujemnych, natomiast $(\Omega, \mathcal{F}, P)$ – przestrzeń probabilistyczną, w której dla każdego $n \in N$ są określone zmienne losowe:

$$\xi_n : \Omega \to S \tag{1}$$

$$\vartheta_n : \Omega \to R_+ \tag{2}$$

Dwuwymiarowy ciąg zmiennych losowych $\{\xi_n, \vartheta_n : n \in N\}$ nazywany jest markowskim procesem odnowy, jeżeli dla każdego $n \in N$, $i, j \in S$, $t \in R_+$:

$$P\{\xi_{n+1} = j, \vartheta_{n+1} < t | \xi_n = i, \xi_{n-1}, \dots, \xi_0, \vartheta_n, \dots \vartheta_0\} = P\{\xi_{n+1} = j, \vartheta_{n+1} < t | \xi_n = i\} \tag{3}$$

oraz

$$P\{\xi_0 = i, \vartheta_0 = 0\} = P\{\xi_0 = i\} \tag{4}$$

Z definicji tej wynika, że markowski proces odnowy jest szczególnym przypadkiem dwuwymiarowego procesu Markowa [14]. Prawdopodobieństwa przejścia tego procesu

zależą wyłącznie od wartości dyskretnej współrzędnej. Markowski proces odnowy $\{\xi_n, \vartheta_n : n \in N\}$ nazywany jest jednorodnym, jeżeli prawdopodobieństwa:

$$P\{\xi_{n+1} = j, \vartheta_{n+1} < t | \xi_n = i\} = Q_{ij}(t) \tag{5}$$

nie zależą od n.
Macierz funkcyjna:

$$Q(t) = \left[Q_{ij}(t)\right], i, j \in S \tag{6}$$

nazywana jest jądrem odnowy. W oparciu o jednorodny markowski proces odnowy definiowany jest proces semi- Markowa [14].
Niech:

$$M(t) = \sup\{m \geq 0 : \tau_m \leq t\} \tag{7}$$

gdzie:

$$\tau_m = \vartheta_0 + \vartheta_1 + \cdots + \vartheta_m \tag{8}$$

Proces stochastyczny $\{M(t): t \in R_+\}$ jest stały w przedziale $[\tau_m, \tau_{m+1})$. Proces stochastyczny $\{X(t): t \in R_+\}$ określony wzorem

$$X(t) = \xi_{M(t)} \tag{9}$$

jest modelem semi –Markowa.

Zdefiniowanie modelowego procesu semi-Markowa wymaga, oprócz określenia jądra procesu również jego rozkładu początkowego [13, 17, 38]. Badany proces eksploatacji pojazdów podzielono na trzy fazy o losowych czasach trwania. Wówczas jądro odnowy procesu semi-Markowa, zgodnie z grafem przejść dozwolonych (rys. 1) przyjmuje postać:

$$Q(t) = \begin{bmatrix} 0 & Q_{12}(t) & Q_{13}(t) \\ Q_{21}(t) & 0 & 0 \\ 0 & Q_{31}(t) & 0 \end{bmatrix} \tag{10}$$

Macierz ta stanowi model zmian wyróżnionych stanów procesu. Niezerowe elementy Q_{ij}(t) macierzy $Q(t)$ są warunkowymi prawdopodobieństwami przejścia procesu ze stanu S_i do stanu S_j, w czasie nie większym niż t, określonymi wg wzoru (11). Zależą one od rozkładu zmiennych losowych, którymi są długości czasu przebywania procesu w wyróżnionych stanach.

$$Q_{ij}(t) = P\big(X(\tau_{m+1}) = j,\ \ \tau_{m+1} - \tau_m \leq t |\ X\ (\tau_m) = \ i)\big) \text{ dla } t \geq 0 \tag{11}$$

gdzie zmienna losowa $\tau_m$oznacza chwilę m-tej zmiany stanu

Rozkład początkowy: $p_i(0), i \in S\ =\ \{1, 2, 3\}$ przyjęto w postaci:

$$p_i(0) = \begin{cases} 1, & \text{gdy } i = 1 \\ 0, & \text{gdy } i \neq 1 \end{cases} \tag{12}$$

gdzie:

$$p_i(0) = P\{X(0) = i\}, \quad i = 1, 2, 3 \tag{13}$$

Elementy te pozwalają wyznaczyć poszukiwane parametry probabilistyczne procesu eksploatacyjnego. Dla modelu semi-Markowa istotne są prawdopodobieństwa przejścia, zdefiniowane jako prawdopodobieństwa warunkowe [15]:

$$P_{ij}(t) = P\{X(t) = j | X(0) = i\}, i, j \in S \tag{14}$$

$P_{ij}(t)$ są to prawdopodobieństwa przejścia ze stanu S_i do stanu S_j w chwili t. Obliczono je na postawie rzeczywistych relacji międzystanowych, wg wzoru (15).

$$p_{ij} = \frac{n_{ij}}{\sum_{k \in S} n_{ik}} \tag{15}$$

gdzie:
n_{ij} – liczba przejść ze stanu S_i do stanu S_j,
$\sum_{k \in S} n_{ik}$ – liczba wszystkich przejść (wyjść) ze stanu S_i,

Rozkład prawdopodobieństwa zmian wyróżnionych stanów eksploatacyjnych (w jednym kroku), przy założeniu, że każdemu łukowi grafu odwzorowania procesu eksploatacji (rys. 1), łączącemu dwa stany procesu, odpowiada wartość prawdopodobieństwa p_{ij} zawiera tab. 1.

Tabela 1. Macierz prawdopodobieństw przejść p_{ij}

p_{ij}	S_1	S_2	S_3
S_1	0	0,8	0,2
S_2	1	0	0
S_3	0	1	0

Obliczone wartości prawdopodobieństw przejść dotyczą zbiorów stanów, a nie okresu czasu. Np. $p_{13} = 0{,}2$ oznacza, że wśród wszystkich wyjść ze stanu S_1 przejścia ze stanu S_1 do S_2 stanowią 20%.

3.2 Własności graniczne

Ważną rolę w badaniu procesu eksploatacji samochodów modelowanych łańcuchem Markowa pełnią jego własności graniczne [13, 20], a szczególnie granice prawdopodobieństw $p_j(n)$ oraz $p_{ij}(n)$ przy $n \to \infty$, które opisują zachowanie procesu po długim czasie [13, 36]. Istotnym pojęciem w tej kwestii jest rozkład stacjonarny jednorodnego łańcucha Markowa, opisywany wektorem Π [14]:

$$\Pi = [\pi_1, \pi_2, \pi_3] \tag{16}$$

takim, że:

$$\Pi = \Pi \mathrm{P} \tag{17}$$

gdzie

$$\mathrm{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \tag{18}$$

oraz

$$\sum_{j=1}^{3} \pi_j = 1 \tag{19}$$

oznacza to, że jeżeli łańcuch w pewnej chwili m osiągnie rozkład stacjonarny, to dla każdej kolejnej chwili n, większej od m, rozkład bezwarunkowy pozostanie taki sam.
Dla badanego procesu istnieją granice:

$$\lim_{n\to\infty} p_{ij}(n) = \pi_j \qquad i,j = 1,2,3 \tag{20}$$

gdzie:
$p_{ij}(n)$ – prawdopodobieństwo przejścia ze stanu S_i do stanu S_j w n krokach.

Obliczona macierz prawdopodobieństw zmian stanów eksploatacyjnych włożonego w proces łańcucha Markowa (tab. 1), umożliwiła wyznaczenie prawdopodobieństw stacjonarnych π_j, zgodnie z układem równań (17).

Dla badanego procesu, dla modelu 3-stanowego, oszacowanie prawdopodobieństw stacjonarnych π_j wymagało rozwiązania równania macierzowego:

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix}^T \cdot \begin{bmatrix} 0 & p_{12} & p_{13} \\ p_{21} & 0 & 0 \\ 0 & p_{32} & 0 \end{bmatrix} = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix}^T \tag{21}$$

z warunkiem normalizacji:

$$\pi_1 + \pi_2 + \pi_3 = 1 \tag{22}$$

co jest równoważne następującemu układowi równań:

$$\begin{cases} \pi_2 \cdot p_{21} = \pi_1 \\ \pi_1 \cdot p_{12} + \pi_3 \cdot p_{32} = \pi_2 \\ \pi_1 \cdot p_{13} = \pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \tag{23}$$

Po podstawieniu wartości prawdopodobieństw przejść (tab. 1) otrzymujemy:

$$\begin{cases} \pi_2 = \pi_1 \\ 0{,}8\,\pi_1 + \pi_3 = \pi_2 \\ 0{,}2\,\pi_1 = \pi_3 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \tag{24}$$

Rozwiązanie układu równań przedstawia tab. 2.

Tabela 2. Prawdopodobieństwa stacjonarne π_j wyróżnionych stanów eksploatacyjnych

	S_1	S_2	S_3
π_i	0,455	0,455	0,09
π_i [%]	45,5	45,5	9

W następnej kolejności, na podstawie grafu skierowanego (rys. 1), określającego prawdopodobieństwa przejść stanów łańcucha Markowa (tab. 1), oraz na podstawie

empirycznych czasów t_{ij} trwania poszczególnych stanów dokonano estymacji warunkowych wartości oczekiwanych $E(T_{ij})$ czasów trwania stanów procesu $X(t)$ na podstawie estymatora określonego wzorem (24)

$$\widehat{E(T_{ij})} = \bar{T}_{ij} = \frac{t_{ij}}{\sum_{j \in S} t_{ij}} \tag{25}$$

Macierz $\bar{T} = [\bar{T}_{ij}],\ i,j = 1,2,3$ oszacowanych warunkowych wartości oczekiwanych czasów T_{ij} przedstawiono w tab. 3.

Tabela 3. Oszacowane wartości oczekiwane warunkowych czasów T_{ij}

$\bar{T}_{ij}[minuty]$	S_1	S_2	S_3
S_1		844	845
S_2	479		
S_3		388	

Znajomość elementów macierzy P i $\bar{T}$ pozwala na oszacowanie wartości oczekiwanych $ET_i, i = 1,2,3$ bezwarunkowych czasów trwania poszczególnych stanów procesu, wg zależności:

$$\widehat{ET_i} = \bar{T}_i = \sum_{j=1}^{3} p_{ij} \cdot \bar{T}_{ij} \tag{26}$$

Dla badanego 3-stanowego procesu użytkowania pojazdów problem oszacowania wartości oczekiwanych bezwarunkowych czasów trwania poszczególnych stanów procesu sprowadził się do rozwiązania następującego układu równań:

$$\begin{cases} \bar{T}_1 = p_{12} \cdot \bar{T}_{12} + p_{13} \cdot \bar{T}_{13} \\ \bar{T}_2 = p_{21} \cdot \bar{T}_{21} \\ \bar{T}_3 = p_{32} \cdot \bar{T}_{32} \end{cases} \tag{27}$$

Oszacowane wartości bezwarunkowych czasów $\bar{T}_i$ przedstawiono w tab. 4.

Tabela 4. Bezwarunkowe czasy $\bar{T}_i$ [minuty] przebywania procesu w 3 stanach eksploatacyjnych

stan	$\bar{T}_i$ [minuty]
1	844,2
2	479
3	388

Zmienne losowe $T_i, i = 1,2,3$ mają skończone, dodatnie wartości oczekiwane. Pozwala to na wyznaczenie rozkładu granicznego procesu semi-Markowa. W oparciu o rozkład stacjonarny włożonego łańcucha Markowa (tab. 2) oraz oszacowane wartości oczekiwane czasów trwania procesu (tab.4) estymowano prawdopodobieństwa graniczne, zgodnie ze wzorem (28) [20].

$$P_i = \frac{\pi_i \cdot \bar{T}_i}{\sum_{k \in S} \pi_k \cdot \bar{T}_k}, \qquad i = 1,2,3 \tag{28}$$

Obliczony, graniczny rozkład prawdopodobieństwa stanów procesu semi-Markowa, przedstawiono w tab. 5.

Tabela 5. Rozkład prawdopodobieństw granicznych P_i

Rozkład	P_1	P_2	P_3
prawdopodobieństwa	0,6026	0,3419	0,0555
procentowy	60	34	6

Wartości P_i są granicznymi prawdopodobieństwami określającymi, że w długim okresie eksploatacji ($t \to \infty$) pojazd będzie przebywał w danym stanie eksploatacyjnym.

Największe wartości osiągnięto dla stanu użytkowanie (60%), co jest bardzo dobrym wynikiem. Postój użytkowy osiąga graniczną wartość wynoszącą 34%, co również jest zadowalającym rezultatem i świadczy z jednej strony o dużej gotowości badanych pojazdów, a z drugiej o znacznej rezerwie, która jednak w przypadku struktur działających w sposób nieprzewidziany, interwencyjny wydaje się być racjonalną. W stanie napraw pojazdy przebywają granicznie jedynie z 5,5% prawdopodobieństwem.

Współczynnik gotowości technicznej K jest sumą odpowiednich prawdopodobieństw stanów niezawodnościowych. Dla zaproponowanego modelu eksploatacji pojazdów stany zdatności stanowią stan S_1 oraz S_2, natomiast stan S_3 jest stanem niezdatności. Stąd gotowość badanych pojazdów można obliczyć jako sumę prawdopodobieństw granicznych stanów S_1 i S_2:

$$K = P_1 + P_2 \tag{29}$$

Obliczony współczynnik gotowości wynosi $K = 94{,}45$ i oznacza, że niemal 95% czasu pojazdy badanej grupy pojazdów pozostają w stanie gotowości technicznej.

3.3 Czas pierwszego przejścia procesu eksploatacji pojazdu do podzbioru stanów (czas bezawaryjnej pracy)

Kolejną, ważną charakterystyką opisującą procesy eksploatacji pojazdów, jest czas pierwszego przejścia rozpatrywanego procesu do wyodrębnionego stanu lub zbioru stanów $\{A\}$ [18]. Na podstawie zidentyfikowania rozkładu tego czasu i jego parametrów można wyznaczyć prawdopodobieństwo przebywania pojazdów w określonym stanie lub zbiorze stanów [20, 37]. Funkcja postaci:

$$\Phi_{iA}(t) = P(\Theta_A \leq t | X(0) = i), t \geq 0 \tag{30}$$

jest dystrybuantą rozkładu zmiennej losowej $\Theta_A = \tau_{\Delta_A}$, która oznacza czas upływający od chwili przyjęcia przez proces semi-Markowa wartości $i \in A'$ do chwili, w której proces przyjmie jakąkolwiek wartość z podzbioru stanów A, gdzie $A \subset S$ oraz $A' = S - A$.
natomiast:

$$\Delta_A = \min\{n \in N: X(\tau_n) \in A\} \tag{31}$$

Dla regularnych procesów semi-Markowa, w których podzbiór A jest silnie osiągalny z każdego stanu należącego do A', zmienne losowe T_{ij} mają skończone i dodatnie wartości oczekiwane $E(T_{ij})$, istnieją wartości oczekiwane $E(\Theta_{A'})$ i są one jedynymi rozwiązaniami układu równań [13, 20]:

$$(I - P_{A'})\overline{\Theta}_{A'} = T_{A'} \tag{32}$$

gdzie:
$P_{A'}$ - macierz prawdopodobieństw przejść w zbiorze A'
$\overline{\Theta}_{A'}$ - jądro procesu określone w zbiorze A'
$T_{A'}$- zmienne losowe bezwarunkowych czasów przebywania procesu w zbiorze stanów A'
Ponieważ w rozważanym procesie zadanie transportowe zostanie wykonane, jeżeli nie nastąpi awaria środka transportu, to rozkład czasu wykonania zadania (bezawaryjnej pracy systemu) można znaleźć redukując pierwotny model o stan S_3 - naprawa. Wówczas podzbiór stanów $A' = \{S_1, S_2\}$, natomiast podzbiór stanów $A = \{S_3\}$, a elementy równania (32) mają postać:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \overline{\Theta}_{A'} = \begin{bmatrix} \overline{\Theta}_{1\,3} \\ \overline{\Theta}_{2\,3} \end{bmatrix}, \qquad \overline{T}_{A'} = \begin{bmatrix} E(T_1) \\ E(T_2) \end{bmatrix}, \tag{33}$$

$$P_{A'}(s) = \begin{bmatrix} 0 & p_{12} \\ p_{2\,1} & 0 \end{bmatrix} \tag{34}$$

gdzie:

$$P_{A'} = [p_{ik}] \quad i, k \in A' \tag{35}$$

jest podmacierzą macierzy P_{ij} (tab. 1). Zmienna losowa Θ_{ij} oznacza czas, który upłynął od chwili początkowej do chwili, w której po raz pierwszy zostanie osiągnięty stan naprawy, pod warunkiem, że w chwili uważanej za początkową rozpoczął się jeden ze stanów ze zbioru A'. Oznacza zatem czas bezawaryjnej eksploatacji systemu. Dla analizowanego modelu semi-Markowa, równanie macierzowe (32) jest postaci:

$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & p_{12} \\ p_{2\,1} & 0 \end{bmatrix}\right) \cdot \begin{bmatrix} \overline{\Theta}_{1\,3} \\ \overline{\Theta}_{2\,3} \end{bmatrix} = \begin{bmatrix} E(T_1) \\ E(T_2) \end{bmatrix} \tag{36}$$

Po podstawieniu odpowiednich wartości z tab.1 i tab.4 otrzymujemy :

$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0{,}8 \\ 1 & 0 \end{bmatrix}\right) \cdot \begin{bmatrix} \overline{\Theta}_{1\,3} \\ \overline{\Theta}_{2\,3} \end{bmatrix} = \begin{bmatrix} 844{,}2 \\ 479 \end{bmatrix} \tag{37}$$

co sprowadza się do rozwiązania układu równań:

$$\begin{cases} \overline{\Theta}_{1\,3} - 0{,}8\,\overline{\Theta}_{2\,3} = 844{,}2 \\ -\overline{\Theta}_{1\,3} + \overline{\Theta}_{2\,3} = 479 \end{cases} \tag{38}$$

Wyniki obliczeń powyższego układu równań przedstawia tab. 6.

Tabela 6. Wartości elementów macierzy $\overline{\Theta}$ czasu pierwszego przejścia dla ogółu pojazdów

$\overline{\Theta}$	$[min]$	$[h]$
$\overline{\Theta}_{1\ 3}$	6137	102,3
$\overline{\Theta}_{2\,3}$	6616	110,3

Jeżeli rozkładem początkowym procesu eksploatacji jest wektor:

$$p = [p_1, p_2, p_3] \tag{39}$$

który w badanym procesie, zgodnie z pierwotnym założeniem (12) jest postaci:

$$p = [1, 0, 0] \tag{40}$$

wówczas pierwszy wiersz jednokolumnowej macierzy, stanowiącej rozwiązanie tego równania, jest wartością oczekiwaną czasu wykonania zadania, która w tym przypadku wynosi ponad 102 godziny.

Możliwe jest również wyznaczenie rozkładu czasu poprawnej pracy obiektu. Korzystając z informacji, że prawdopodobieństwa przejścia $P_{ij}(t)$, zdefiniowane jako prawdopodobieństwa warunkowe [20]:

$$P_{ij}(t) = P\{X(t) = j|X(0) = i\}, \qquad i,j \in S \tag{41}$$

spełniają równania Fellera:

$$P_{ij}(t) = \delta_{ij}[1 - G_i(t)] + \sum_{k\in S}\int_0^t P_{kj}(t-x)dQ_{ik}(x)\,, \qquad i,j \in S \tag{42}$$

można znaleźć rozwiązanie tego układu posługując się przekształceniem Laplace'a – Stieltiesa:

$$\tilde{p}_{ij}(s) = \int_0^\infty e^{-st}\,dP_{ij}(t) \tag{43}$$

$$\tilde{q}_{ik}(s) = \int_0^\infty e^{-st}\,d\mathrm{Q}_{ik}(t) \tag{44}$$

$$\tilde{g}_i(s) = \int_0^\infty e^{-st}\,d\mathrm{G}_i(t) \tag{45}$$

gdzie $Q_{ik}(t)$ to jądro procesu odnowy podzbioru stanów A' natomiast $\mathrm{G}_i(t)$ oznacza dystrybuantę zmiennej losowej T_i czasu trwania *i*-tego stanu procesu semi-Markowa, niezależnie od tego do jakiego stanu następuje przejście w chwili τ_{n+1} [13]:

$$G_i(t) = P\{T_i < t\} = P\{\tau_{n+1} - \tau_n < t/X(\tau_n) = 1\}, i \in S \tag{46}$$

Wówczas powyższemu układowi równań całkowych odpowiada układ równań algebraicznych, o niewiadomych transformatach $p_{ij}(s),\ i,j \in S$:

$$\tilde{p}_{ij}(s) = \delta_{ij}\left[\frac{1-\tilde{g}_i(s)}{s}\right] + \sum_{k\in S}\tilde{q}_{ik}(s)\tilde{p}_{kj}(s)\,, \qquad i,j \in S \tag{47}$$

układ ten w notacji macierzowej ma postać:

$$\tilde{P}(s) = \frac{1}{s}[I - \tilde{q}(s)]^{-1}\,[1 - \tilde{g}(s)] \tag{48}$$

Po rozwiązaniu otrzymuje się macierz transformat. Ponieważ stanem początkowym jest stan S_1, więc pierwszy wiersz jest jednocześnie jednowymiarowym rozkładem procesu.
Dla badanego systemu:

$$Q(t) = \begin{bmatrix} 0 & Q_{12}(t) \\ Q_{21}(t) & 0 \end{bmatrix} \tag{49}$$

gdzie $Q_{12}(t)$ oraz $Q_{21}(t)$ są dystrybuantami estymowanych rozkładów Gamma:

$$Q_{12}(t) = \frac{e^{-\frac{t}{\beta_1}}t^{-1+\alpha_1}\beta^{-\alpha_1}}{\Gamma[\alpha_1]}, \quad t > 0 \tag{50}$$

$$Q_{21}(t) = \frac{e^{-\frac{t}{\beta_2}} t^{-1+\alpha_2} \beta^{-\alpha_2}}{\Gamma[\alpha_2]} \quad t > 0 \tag{51}$$

oraz

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} \tag{52}$$

ponieważ dla rozkładu Gamma transformata Laplace'a – Stieltjesa jest postaci:

$$\tilde{f}(s) = \left(\frac{\beta}{\beta + s}\right)^\alpha \tag{53}$$

stąd elementy równania (48) mają postać:

$$\frac{1}{s}[I - \tilde{q}(s)]^{-1} = \begin{bmatrix} \frac{0.1}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} & \frac{0.1\,\beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} \\ \frac{0.1\,\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} & \frac{0.1}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} \end{bmatrix} \tag{54}$$

oraz

$$[1 - \tilde{g}(s)] = \begin{bmatrix} 1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1} & 1 \\ 1 & 1 - \beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1} \end{bmatrix} \tag{55}$$

Rozwiązaniem jest macierz, której elementy pierwszego wiersza wynoszą:

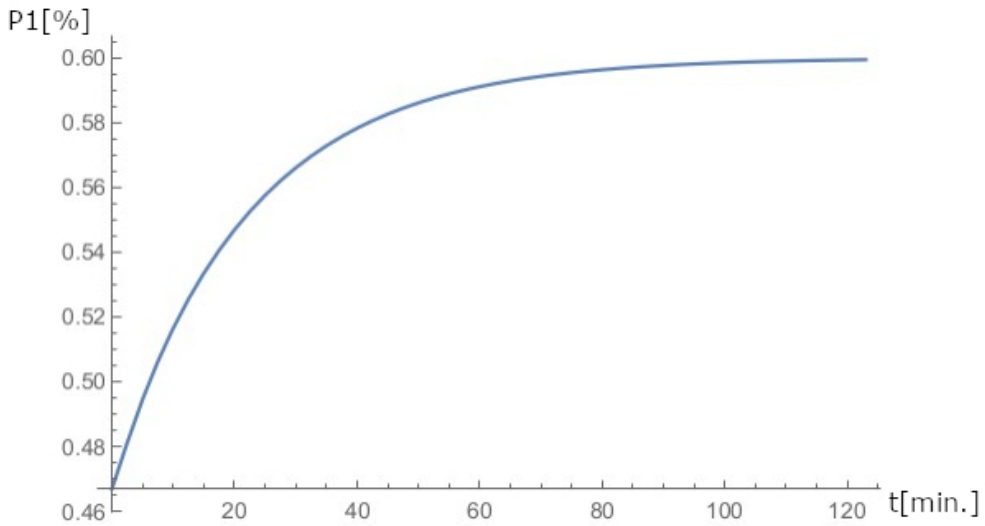
$$\tilde{P}_1(s) = \frac{0.1\beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} + \frac{0.1\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\right)}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} \tag{56}$$

$$\tilde{P}_2(s) = \frac{0.1}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} + \frac{0.1\beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\left(1 - \beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)}{s\left(1 - \beta_1^{\alpha_1}(s+\beta_1)^{-\alpha_1}\beta_2^{\alpha_2}(s+\beta_2)^{-\alpha_1}\right)} \tag{57}$$

Po obliczeniu transformat odwrotnych otrzymuje się rozkład graniczny intensywności użytkowania obiektu. Dla stanu S_1 otrzymujemy funkcję postaci:

$$P_1(t) = 0.0026857\, e^{-0.66956t} - 0.006151\, e^{-0.23043t} - 0.129546\, e^{-0.0446957t} + 0.6 \tag{58}$$

Wykres tej funkcji przedstawia rys. 3



Rysunek 3. Wykres funkcji $P_1(t)$

Funkcja stabilizuje się w czasie około 120 minut, a w granicy, dla $t \to \infty$, dąży do wcześniej obliczonej wartości granicznej procesu semi-Markowa wynoszącej $P_1 = 60\%$.

4. Wnioski

Zastosowanie procesów semi -Markowa pozwala na wyznaczenie granicznego współczynnika gotowości oraz analizę czasów przebywania pojazdów specjalnych w wyróżnionych stanach eksploatacyjnych. Umożliwia także obiektywną ocenę intensywności użytkowania pojazdu i czasu jego bezawaryjnej pracy. Analizując czynniki gotowości można poszukiwać optymalnych algorytmów użytkowania i obsługiwania pojazdów, a także analizować jakość doboru floty pojazdów.

Słuszność powyższych założeń potwierdziły zrealizowane badania. Zaproponowany model semi-Markowa umożliwił diagnostykę systemu eksploatacji radiowozów policyjnych wskazując, że charakteryzuje się on zadowalającym poziomem prawdopodobieństwa przebywania pojazdów w stanie użytkowania ($P_1 = 0{,}6$) oraz w stanie postoju użytkowego ($P_2 = 0{,}34$). Prognozowany współczynnik gotowości technicznej wyniósł $K = 95\%$.

Wykazano zatem skuteczność zastosowania procesów semi-Markowa do modelowania gotowości systemów eksploatacji pojazdów specjalnych. Model trzystanowy wyróżniający stan użytkowania pojazdu i stan postoju użytkowego oraz stan naprawy (obsługiwania technicznego) okazał się uzasadniony. W tym przypadku nie było konieczne tworzenie rozbudowanych, wielostanowych struktur modelu procesu eksploatacji wymagających zaawansowanych programów obliczeniowych. Zaprezentowany, trzystanowy model jest możliwy do rozbudowy w sytuacji, kiedy konieczna byłaby pogłębiona analiza wybranych aspektów gotowości systemu.

Bibliografia

1. Andrzejczak K, Młyńczak M, Selech J. Poisson-distributed failures in the predicting of the cost of corrective maintenance. Eksploatacja I Niezawodnosc - Maintenance and Reliability 2018; 20(4): 602-609, https://doi.org/10.17531/ein.2018.4.11.
2. Bain L.J, Engelhardt M. Introduction to Probability and Mathematical Statistics. Second Edition. California: Cengage Learning, 2000.

3. Becker L.R, Zaloshnja E, Levick N, Guohua L, Miller T. R. Relative risk of injury and death in ambulances and other emergency vehicles. Accident Analysis & Prevention 2003; 35(6): 941-948, https://doi.org/10.1016/S0001-4575(02)00102-1.
4. Behm G.W, Huber W.B, Noll A.J, Pelaez R. A Method and system for safe emergency vehicle operation using route calculation. United States Patent US8842021B2, 2014.
5. Cheng Q, Sun B, Zhao Y, Gu P. A method to analyze the machining accuracy reliability sensitivity of machine tools based on Fast Markov Chain simulation. Podejście do analizy czułości niezawodnościowej dokładności obrabiarek oparte na symulacji metodą szybkich łańcuchów Markowa. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (4): 552-564, https://doi.org/10.17531/ein.2016.4.10.
6. Chu H. C. Risk factors for the severity of injury incurred in crashes involving on-duty police cars. Traffic injury prevention 2016, (5)17: 495-501, https://doi.org/10.1080/15389588.2015.1109082.
7. Dekker R, Nicolai R.P, Kallenberg L.C.M, Maintenance and Markov decision models. In Wiley StatsRef: Statistics Reference Online (eds Balakrishnan N, Colton T, Everitt B, Piegorsch W, Ruggeri F, Teugels J.L.). John Wiley & Sons, 2014, https://doi.org/10.1002/9781118445112.stat03960.
8. Dinc S, Dinc I. Evaluation of Unsupervised Classification on Police Patrol Zone Design Problem. SoutheastCon 2018. St Petersburg, 2018: 1-7, https://doi.org/10.1109/SECON.2018.8478908.
9. Dong W, Liu S, Yang X, Wang H, Fang Z. Balancing reliability and maintenance cost rate of multi-state components with fault interval omission. Eksploatacja I Niezawodnosc - Maintenance and Reliability 2019; 21(1): 37-45, https://doi.org/10.17531/ein.2019.1.5.
10. Elliott T, Payne A, Atkison T, Smith R. Algorithms in Law Enforcement: Toward Optimal Patrol and Deployment Algorithms. Proceedings of the 2018 International Conference on Information and Knowledge Engineering IKE'18. Las Vegas, 2018: 93-99.
11. Ge H, Tomasevicz C.L, Asgarpoor S. Optimum Maintenance Policy with Inspection by Semi-Markov Decision Processes. 39th North American Power Symposium, Las Cruces, 2007: 541-546, https://doi.org/10.1109/NAPS.2007.4402363.
12. Girtler J, Ślęzak M. Application of the theory of semi-Markov processes to the development of a reliability model of an automotive vehicle. Archiwum Motoryzacji 2012; 2: 15-27, https://doi.org/10.5604/1234754X.1066721.
13. Grabski F. Semi-Markov Processes. Applications in System Reliability and Maintenance. Elsevier, 2015, https://doi.org/10.1016/B978-0-12-800518-7.00004-1.
14. Grabski F. Teoria semi-Markowskich procesów eksploatacji obiektów technicznych. - The theory of semi-Markov processes of technical object exploitation Gdynia: Zeszyty Naukowe Wyższej Szkoły Marynarki Wojennej 75A, 1982.
15. Hong W, Zhou K. A note on the passage time of finite-state Markov chains. Communications in Statistics - Theory and Methods 2017; 46(1): 438-445, https://doi.org/10.1080/03610926.2014.995825.
16. Hu L, Su P, Peng R, Zhang Z. Fuzzy Availability Assessment for Discrete Time Multi-State System under Minor Failures and Repairs by Using Fuzzy Lz-transform. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19 (2): 179-190, https://doi.org/10.17531ein.2017.2.5.
17. Huang X.X, Zou X.L, Guo X.P. A minimization problem of the risk probability in first passage semi-Markov decision processes with loss rates. Science China Mathematics 2015, 58: 1923 - 1938, https://doi.org/10.1007/s11425-015-5029-x.

18. Hunter J.J. The computation of the mean first passage times for Markov chains. Linear Algebra and its Applications 2018; 549: 100-122, https://doi.org/10.1016/j.laa.2018.03.010.
19. Iscioglu F, Kocak A. Dynamic reliability analysis of a multi-state manufacturing system. Eksploatacja i Niezawodnoscć - Maintenance and Reliability 2019; 21 (3): 451-459, https://doi.org/10.17531/ein.2019.3.11.
20. Jaźwiński J, Grabski F. Niektóre problemy modelowania systemów transportowych - Selected problems of transport system modelling. Radom: Instytut Technologii Eksploatacji, 2003.
21. Kaczor G. Modelowanie i ocena niezawodności systemu transportu intermodalnego - Modelling and assessment of the reliability of the intermodal transport system. Logistyka 2015; 3: 2047-2054.
22. Kolesar P.J, Rider K.L, Crabill T.B, Walker W.E. A Queuing-Linear Programming Approach to Scheduling Police Patrol Cars. Operations Research 1975; 23(6):1045-1062, https://doi.org/10.1287/opre.23.6.1045.
23. Landowski B, Muślewski Ł, Knopik L, Bojar P. Semi-Markov model of quality state changes of a selected transport system. Journal of KONES 2017; 24(4): 141-148.
24. Lu J-M, Lundteigen M.A, Liu Y, Wu X-Y. Flexible truncation method for the reliability assessment of phased mission systems with repairable components. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (2): 229-236, https://doi.org/10.17531/ein.2016.2.10.
25. Lundälv J, Philipson Ch, Sarre R. How do we reduce the risk of deaths and injuries from incidents involving police cars? Understanding injury prevention in the Swedish context. Police Practice and Research 2010; 11(5): 437-450, https://doi.org/10.1080/15614263.2010.497333.
26. Lyons H.W. Integrated warning light and rear-view mirror. United States Patent 5851064, 1998.
27. Michaelson E.B. Bulletproof blanket for use with law enforcement vehicles such as police cars. United States Patent 6161462, 2000.
28. Migawa K. Availability control for means of transport in decisive semi-Markov models of exploitation process. Archives of Transport 2012; 4(24): 497-508, https://doi.org/10.2478/v10174-012-0030-4.
29. Młyńczak M. Metodyka badań eksploatacyjnych obiektów mechanicznych - Methodology of exploitation tests of mechanical objects. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, 2012.
30. Muślewski Ł. Control Method for Transport System Operational Quality. Journal of KONES 2009; 3(16): 275-282.
31. Restel F. The Markov reliability and safety model of the railway transportation system. Safety and Reliability: Methodology and Applications Proceeding of the European Safety and Reliability Conference. London, 2014: 303-311, https://doi.org/10.1201/b17399-46.
32. Świderski A. Inżynieria jakości w wybranych obszarach transportu - Quality engineering in selected areas of transport. Warszawa: Instytut Transportu Samochodowego (Motor Transport Institute), Warszawa 2018.
33. Szawłowski S. Analiza wpływu systemu obsług na gotowość techniczną śmigłowca pokładowego SH-2G - Analysis of the impact of the maintenance system on the technical readiness of the SH-2G ship-based helicopter. Prace Instytutu Lotnictwa 2008; 3-4 (194-195): 326-331.
34. Thomas O.S, Sobanjo J.O. Semi-Markov Decision Process: A Decision Tool for Transportation Infrastructure Management Systems. International Conference on

Transportation and Development: Projects and Practices for Prosperity 2016: 384 - 396, https://doi.org/10.1061/9780784479926.036.
35. Woropay M, Żurek J, Migawa K. Model of assessment and shaping of operational readiness of the maintenance subsystem in the transport system. Radom: Instytut Technologii Eksploatacji, 2003.
36. Wu X, Zhang J. Finite approximation of the first passage models for discrete-time Markov decision processes with varying discount factors. Discrete Event Dynamic Systems 2016; 26(4): 669 - 683, https://doi.org/10.1007/s10626-014-0209-3.
37. Wu X, Zou X, Guo X. First passage Markov decision processes with constraints and varying discount factors. Frontiers of Mathematics in China 2015; 10(4): 1005-1023, https://doi.org/10.1007/s11464-015-0479-6.
38. Xie W, Hong Y, Trivedi K. Analysis of a two-level software rejuvenation policy. Reliability Engineering & System Safety 2005; 87: 13-22, https://doi.org/10.1016/j.ress.2004.02.011.
39. Żurek J, Tomaszewska J. Analysis of the exploitation system from the standpoint of readiness, Prace Naukowe Politechniki Warszawskiej, Warszawa 2016; 114: 471 -477.

dr hab. inż. Grzegorz Żywica

dr inż. Tomasz Z. Kaczmarczyk

Department of Turbine Dynamics and Diagnostics
Institute of Fluid Flow Machinery, Polish Academy of Sciences
Fiszera 14 str., 80-231 Gdansk, Poland
E-mail: gzywica@imp.gda.pl, tkaczmarczyk@imp.gda.pl

Eksperymentalna ocena właściwości dynamicznych mikroturbiny energetycznej w obecności defektów układu wirującego

Experimental evaluation of dynamic properties of the energy microturbine with defects in the rotating system

Słowa kluczowe: *mikroturbiny, wysokoobrotowe wirniki, uszkodzenia układu wirującego, drgania maszyn wirnikowych, dynamika wirników*

Keywords: *microturbines, high-speed rotors, damage in a rotating system, vibrations of turbomachines, rotor dynamics*

Streszczenie: We współczesnych systemach energetycznych coraz częściej do wytwarzania energii elektrycznej stosowane są różnego typu mikroturbiny. Charakterystyczną cechą takich maszyn są wysokoobrotowe wirniki, których prędkości obrotowe mogą przekraczać nawet 100 000 obr/min. Praca wirnika w takich warunkach wymaga zastosowania specjalnych rozwiązań konstrukcyjnych i bardzo dużej precyzji wykonania, a podczas eksploatacji zachowania odpowiednich procedur przy rozruchu i odstawieniu, a także stosowania dedykowanych systemów diagnostycznych. W niniejszym artykule zostały omówione badania eksperymentalne mikroturbiny parowej o mocy 2,5 kW, pracującej w prototypowym układzie kogeneracyjnym. Wykonane pomiary obejmowały ocenę stanu dynamicznego podczas normalnej pracy maszyny oraz badania jej właściwości dynamicznych w obecności defektów układu wirującego. Uzyskane wyniki pomiarów, w postaci rozkładów częstotliwościowych drgań, pozwalają na zdefiniowanie symptomów diagnostycznych typowych dla różnych defektów, które mogą pojawić się podczas eksploatacji tej klasy maszyn wirnikowych.

Abstract: Today's energy systems increasingly use various types of microturbines to generate electricity. A specific feature of such a type of machines is a high-speed rotor, whose rotational speed can be higher than 100,000 rpm. Failure-free operation of microturbine rotors at high speeds requires both special design and high precision of the manufacturing process. What is more, proper procedures must be followed during run-up and coast-down phases; also, dedicated diagnostic systems have to be used. This article discusses the experimental research conducted on a 2.5 kW vapour microturbine that operated in a prototypical combined heat and power plant. A series of measurements was carried out to evaluate the dynamic performance of the machine during normal operation. After certain defects of the rotating system occurred, there was a need to perform a new series of measurements in order to assess the dynamic properties of the machine. The obtained measurement results, in the form of vibration velocity spectra, allowed to define diagnostic symptoms corresponding to particular defects. Similar diagnostic symptoms can occur during the operation of this class of turbomachines.

1. Wprowadzenie

W nowoczesnych systemach elektroenergetycznych coraz większą rolę odgrywają rozproszone źródła energii, które umożliwiają efektywne wytwarzanie energii cieplnej i elektrycznej w małej skali, w oparciu o lokalne zasoby energetyczne [3]. W zależności od dostępnego źródła energii pierwotnej lub odnawialnej oraz zapotrzebowania energetycznego, do wytwarzania energii elektrycznej w takich układach mogą być stosowane różnego typu mikroturbiny cieplne [21]. Ich cechami charakterystycznymi są: możliwość dostosowania mocy do zapotrzebowania, duża gotowość eksploatacyjna (w przeciwieństwie do turbin parowych dużej mocy, mogą być szybko uruchamiane i odstawiane), małe wymiary gabarytowe, wysokie prędkości obrotowe wirników, duża mobilność oraz stosunkowo niskie koszty inwestycyjne i eksploatacyjne [11]. Te cechy mikroturbin sprawiają, że są one coraz częściej stosowane, przyczyniając się w dużym stopniu do rozwoju małych układów kogeneracyjnych [12].

Pojęcie „mikroturbiny" najczęściej stosowane jest w odniesieniu do turbin o mocy elektrycznej nieprzekraczającej 1 MW. Mikroturbiny można podzielić na kilka grup w zależności od: zasady działania, mocy, rodzaju paliwa, konstrukcji układu przepływowego, czy zastosowania. Główny podział mikroturbin cieplnych rozróżnia mikroturbiny gazowe i parowe. W mikroturbinach gazowych do napędu wirnika wykorzystuje się gazy wylotowe z komory spalania [2, 21]. Jako paliwo można w nich wykorzystać np. gaz ziemny, biogaz lub naftę. W mikroturbinach parowych wirnik napędzany jest sprężoną parą czynnika roboczego, znajdującego się w obiegu zamkniętym, który jest ogrzewany z zewnętrznego źródła ciepła [12]. Źródłem ciepła w takim układzie może być: kocioł na biomasę lub gaz, źródło geotermalne, kolektory słoneczne lub ciepło odpadowe z różnych procesów produkcyjnych [7, 18]. Instalacje z mikroturbinami gazowymi i parowymi najwyższe sprawności i rentowność uzyskują przy pracy kogeneracyjnej, tzn. gdy oprócz energii elektrycznej zagospodarowana jest również energia cieplna. Mikroturbiny parowe pracują przy niższych temperaturach niż mikroturbiny gazowe i umożliwiają wytwarzanie energii elektrycznej ze źródeł niskotemperaturowych. W przeciwieństwie do mikroturbin gazowych nie są one również wrażliwe na jakość paliwa, gdyż są zasilane odseparowanym od komory spalania czynnikiem roboczym. Jako medium robocze mikroturbin parowych stosowana jest najczęściej woda, ale przy niższych temperaturach stosowane są również różne czynniki niskowrzące [6]. Takie obiegi termodynamiczne nazywa się organicznymi obiegami Rankine'a (ang. organic Rankine cycle).

W mikroturbinach energetycznych stosowane są wysokoobrotowe wirniki. Przy wysokich prędkościach obrotowych wału (przekraczających nawet 100 000 obr/min [13]) wymagane jest stosowanie zaawansowanych, niekonwencjonalnych systemów łożyskowania, takich jak łożyska: magnetyczne [15], foliowe [28], gazowe [19] lub smarowane czynnikiem niskowrzącym [14]. Poza bardzo dobrymi właściwościami dynamicznymi wirnika w szerokim zakresie prędkości, systemy łożyskowania muszą zapewniać wysoką trwałość i niezawodność [7], umożliwiając eksploatację mikroturbin bez stałego nadzoru technicznego. Dzięki małym rozmiarom oraz mniejszej mocy mikroturbiny nie wymagają zachowania tak restrykcyjnych procedur rozruchowych, jak w przypadku dużych turbin energetycznych [5]. Ze względu na małe wymiary i dużą sztywność łopatek turbinowych nie występują w nich również problemy dynamiczne, typowe dla maszyn wirnikowych stosowanych w energetyce zawodowej [16]. Mikroturbiny mogą jednak sprawiać inne problemy dynamiczne, wynikające m.in. z: wysokich prędkości obrotowych, szerokiego zakresu prędkości roboczych, małej masy i sztywności korpusów oraz wiotkich i lekkich konstrukcji podpierających. Niekorzystny wpływ na podzespoły mikroturbin mogą mieć również niektóre czynniki niskowrzące stosowane w układach ORC, gdyż mogą one oddziaływać chemicznie na

materiały konstrukcyjne, w szczególności na tworzywa sztuczne i elastomery. Ponieważ czynniki niskowrzące nie są tak popularne jak woda, nie zawsze konstruktor posiada pełną informację o ich kompatybilności chemicznej z różnymi materiałami.

Ponieważ we wcześniejszych latach mikroturbiny energetyczne były rzadko spotykanym rozwiązaniem, a ich produkcją i badaniami zajmowały się głównie duże, światowe koncerny, w literaturze można znaleźć niewiele informacji na temat szczegółów konstrukcyjnych oraz wytycznych dotyczących eksploatacji tego typu maszyn. Dotyczy to również wyników badań właściwości dynamicznych mikroturbin funkcjonujących w różnych warunkach pracy. Tymczasem dla osób zajmujących się eksploatacją i utrzymaniem ruchu maszyn wirnikowych wyniki takich badań dostarczają bardzo cennych informacji, umożliwiających opracowanie odpowiedniego systemu diagnostycznego, ustawienie progów ostrzegawczych i alarmowych, czy określenie okresów międzyprzeglądowych. Wyniki pomiarów drgań maszyn wirnikowych umożliwiają również wprowadzenie zmian w konstrukcji maszyny oraz układzie podpierającym, pozwalających na poprawę ich stanu dynamicznego.

Dość często obiektami badań naukowców i inżynierów zajmujących się budową i eksploatacją maszyn wirnikowych są duże turbiny parowe. W tym przypadku oprócz wyników badań przeprowadzonych na poprawnie funkcjonujących maszynach, można znaleźć również publikacje, w których analizowane są różnego typu defekty. Tego typu badania eksperymentalne pomocniczej turbiny parowej o mocy 15 MW zostały przedstawione w pracy [22]. W omawianym przypadku uszkodzeniu uległ drut tłumiący drgania łopatek turbinowych jednego stopnia części niskoprężnej. W celu wyjaśnienia przyczyn awarii wykonano pomiary drgań oraz eksperymentalną analizę modalną łopatki. Dodatkowa analiza numeryczna umożliwiła wskazanie miejsca na łopatce, w którym wystąpiło spiętrzenie naprężeń, co pozwoliło jednoznacznie wyjaśnić przyczynę problemu. Badania na temat przyczyn wystąpienia uszkodzenia układu łopatkowego części niskoprężnej turbiny parowej o mocy 310 MW zostały omówione również w pracy [27]. Uszkodzenia łopatek w tym przypadku były spowodowane korozją wgłębną. Autorzy skupili się na analizie materiałowej miejsca uszkodzenia, a jako przyczynę uszkodzenia wskazali niewłaściwą obróbkę cieplną łopatek. Badania diagnostyczne wirnika turbiny energetycznej o mocy 60 MW zostały omówione w pracy [1]. Po 10 latach eksploatacji pojawiło się pęknięcie zmęczeniowe w przekroju wału. W celu jego wykrycia autorzy artykułu zastosowali różne techniki diagnostyczne, ale skupiono się głównie na ocenie przełomu zmęczeniowego i analizach materiałowych. W pracy [17] omówione zostały drgania skrętne i ocena zużycia zmęczeniowego wału turbiny parowej o mocy 600 MW po regeneracji. W badaniach uwzględniono również koncentrację naprężeń w miejscu, w którym wystąpiło największe zużycie. W pracy [24] przedstawiona została analiza wpływu zmian temperatury na uszkodzenia zmęczeniowe wirnika turbiny parowej o mocy 1000 MW. Poursaeidi i inni [23] przeprowadzili z kolei analizę wpływu rozkładu temperatury na deformacje i spiętrzenia naprężeń w korpusie turbiny gazowej. Analizę z wykorzystaniem modelu MES przeprowadzili dla trzech poziomów mocy: 82, 87 i 96 MW. Podczas tworzenia modelu wykorzystano wyniki pomiarów temperatury uzyskane na obiekcie rzeczywistym. Wykazano, że pojawienie się pęknięć w obrębie niektórych otworów korpusu ma związek z koncentracją naprężeń cieplnych. Przykład zastosowania metody elementów skończonych do analizy przyczyn powstawania uszkodzeń korpusu silnika turbinowego stosowanego do napędu helikoptera został przedstawiony w artykule [25]. W analizie uwzględniono obciążenia cieplne i mechaniczne. Również w tym przypadku pęknięcia powstawały w miejscach największej koncentracji naprężeń cieplnych. Zauważono, że drgania wirnika mogą przyczynić się do szybszego pojawiania się uszkodzeń korpusu. Wyniki pomiarów drgań dużej maszyny wirnikowej (wentylatora promieniowego) w stanach ustalonych oraz

nieustalonych zostały przedstawione w pracy [4]. Badania wykonano podczas normalnej eksploatacji, co umożliwiło wskazanie potencjalnych źródeł podwyższonego poziomu drgań. Nadmierne drgania były spowodowane między innymi zbyt dużym niewyważeniem układu wirującego oraz niewłaściwą regulacją zaworów wlotowych. W przypadku dużych turbin parowych trudno jest prowadzić badania eksperymentalne w obecności uszkodzeń elementów układu wirującego, gdyż mogłoby to doprowadzić do poważnej w skutkach awarii oraz zagrażałoby życiu i zdrowiu obsługi. W związku z tym trudno jest również znaleźć artykuły, w których przedstawiono by wyniki takich pomiarów. Dlatego do analizy uszkodzeń w takich maszynach powszechnie wykorzystywane są różnego typu modele numeryczne.

Najczęściej publikowane wyniki badań eksperymentalnych mikroturbin energetycznych dotyczą głównie pomiaru mocy i prędkości obrotowej oraz różnych parametrów termodynamicznych. Znacznie mniej uwagi poświęca się pomiarom drgań oraz badaniom różnego typu defektów, które mogą pojawić się podczas eksploatacji. W pracy [10] przedstawiony został przykład systemu diagnostycznego dedykowanego dla układu kogeneracyjnego z mikroturbiną gazową o mocy 95 kW. System umożliwiał wykrywanie defektów układu kogeneracyjnego oraz niesprawności układu pomiarowego w oparciu o pomiary mocy cieplnej i elektrycznej, parametry termodynamiczne na wylocie mikroturbiny oraz prędkość obrotową wirnika (która dochodziła do 70 tys. obr/min). W systemie diagnostycznym nie wykorzystano jednak możliwości, jakie daje pomiar drgań. W artykule [26] przedstawione zostały badania właściwości dynamicznych wysokoobrotowego wirnika (30 tys. obr/min) o złożonej geometrii układu łopatkowego. Został on zaprojektowany do zastosowania na stanowisku służącym do badania mikroturbin gazowych o mocy 100 kW. Do podparcia wirnika użyto ceramicznych łożysk tocznych. Oprócz wyników badań eksperymentalnych przedstawiony został również model numeryczny wirnika, w którym uwzględniono zastępcze współczynniki sztywności łożysk. Model umożliwiał szczegółową analizę właściwości dynamicznych i cieplnych. Stwierdzono, że drgania wirnika przy niższych prędkościach były bardzo małe dzięki wysokiej prędkości krytycznej, co przekładało się również na niezawodną pracę maszyny. Hong i inni przedstawili badania dynamiki wirnika mikroturbiny gazowej o mocy 500W i prędkości nominalnej równej 100 000 obr/min [8]. Obliczenia i badania eksperymentalne zostały wykonane dla różnych poziomów niewyważenia wirnika. Dodatkowo opracowany został model obliczeniowy, który pozwolił na analizę drgań wymuszonych wirnika.

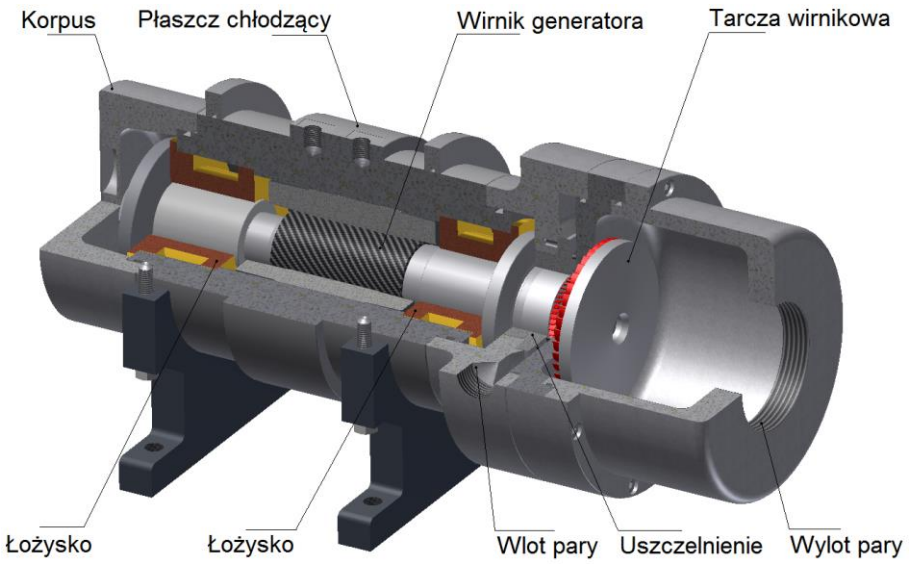
Wyniki badań mikroturbin energetycznych, które zostały dotychczas opublikowane w literaturze, tylko w nielicznych przypadkach uwzględniają pomiary drgań. Autorzy niniejszego artykułu nie znaleźli również publikacji, w których przedstawiono by wyniki pomiarów drgań tego typu maszyn wirnikowych, uzyskane w obecności defektów lub przy niepoprawnej pracy. Tymczasem badania mikroturbin dają znacznie większe możliwości w zakresie prowadzenia czynnych eksperymentów diagnostycznych. Ich przestoje nie wiążą się z dużymi stratami finansowymi, a znacznie mniejsze koszty produkcji umożliwiają wprowadzenie różnych modyfikacji i wcześniejsze przygotowanie części zamiennych. Pomimo, że ich wirniki obracają się z dużymi prędkościami obrotowymi, to ich niewielkie masy i małe wymiary umożliwiają zastosowanie całkowicie pewnych zabezpieczeń, pozwalających na bezpieczne prowadzenie takich eksperymentów. W porównaniu do dużych turbin energetycznych, mikroturbiny dają więc znacznie większe możliwości w zakresie prowadzenia badań w obecności różnych defektów.

W dalszej części artykułu omówione zostały badania diagnostyczne mikroturbiny parowej ORC o mocy elektrycznej do 2,5 kW. Pomiary drgań zostały wykonane w obecności różnego typu defektów układu wirującego, które pojawiły się w trakcie eksploatacji. Badania wibrodiagnostyczne tej mikroturbiny były wcześniej prezentowane w literaturze [9], ale w trakcie wcześniejszych pomiarów nie zostały wykryte żadne problemy dynamiczne, ani

symptomy świadczące o możliwości pojawienia się defektów. Dlatego wcześniejsze badania należy traktować jako bazowe, do których można się odnieść w przypadku zmian charakterystyk drganiowych. Uszkodzenia, które pojawiły się po pewnym czasie eksploatacji mikroturbiny pozwoliły na uzyskanie nowych charakterystyk, których analiza umożliwia wykrycie różnego typu defektów na wczesnym etapie ich rozwoju. Daje to duże możliwości w zakresie właściwej obsługi i zapobiegania awariom podczas pracy.

2. Charakterystyka badanej mikroturbiny

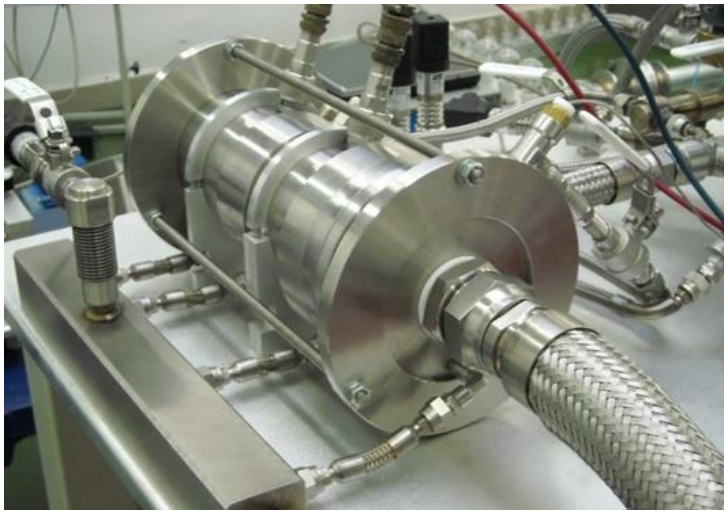
Badana mikroturbina została opracowana jako urządzenie służące do wytwarzania energii elektrycznej w małym układzie kogeneracyjnym ORC [29]. Jest to maszyna prototypowa, która powstała w ramach projektu POIG.01.01.02-00-016/08, w wyniku współpracy Instytutu Maszyn Przepływowych PAN w Gdańsku z Instytutem Maszyn Przepływowych Politechniki Łódzkiej. Maksymalna moc cieplna i elektryczna układu kogeneracyjnego została dopasowana do przeciętnego zapotrzebowania domów jednorodzinnych. Przy mocy cieplnej kotła dochodzącej do 25 kW, za pomocą mikroturbiny można uzyskać do 2,5 kW mocy elektrycznej. Ze względu na przyszłe zastosowanie mikroturbina wraz z generatorem charakteryzuje się bardzo małymi wymiarami. Długość korpusu wynosi ok. 350 mm, a jego średnica zewnętrzna nie przekracza 200 mm. Przekrój mikroturbiny wraz z podstawowymi częściami został przedstawiony na rys. 1. Na rysunku tym widać sposób umieszczenia wirnika w korpusie oraz system łożyskowania. Do podparcia wału zastosowano dwa poprzeczno-wzdłużne łożyska gazowe, stale zasilane parą czynnika niskowrzącego o oznaczeniu HFE-7100 (tego samego czynnika, który zasila układ przepływowy mikroturbiny). Dzięki zastosowaniu łożysk gazowych, do ich smarowania nie stosuje się oleju, a wirnik może uzyskiwać bardzo wysokie prędkości obrotowe przy minimalnych stratach tarcia w węzłach łożyskowych. Ponieważ czynnik smarny łożysk jest pobierany z obiegu ORC, nie są potrzebne żadne dodatkowe układy smarowania. Nie ma również zagrożenia wystąpienia przecieku oleju do czynnika niskowrzącego. Tego typu rozwiązania konstrukcyjne mikroturbin nazywa się bezolejowymi.



Rys. 1. Przekrój częściowy mikroturbiny z oznaczonymi głównymi częściami

Układ przepływowy mikroturbiny składa się z czterech stopni turbinowych, przy czym przepływ pary przez dwa stopnie odbywa się w kierunku dośrodkowym, a przez pozostałe dwa stopnie w kierunku odśrodkowym. Dzięki takiemu rozwiązaniu

minimalizowana jest siła osiowa, działająca na łożyska wzdłużne. Ciśnienie pary czynnika niskowrzącego na wlocie do mikroturbiny wynosi ok. 11 bar, a jego temperatura ok. 180°C. Nominalna prędkość obrotowa wirnika wynosi 24 000 obr/min, ale maszyna ta jest również przystosowana do pracy ciągłej przy niższych prędkościach. Pomiędzy czopami łożyskowymi na wale mikroturbiny umieszczona została tuleja generatora synchronicznego. Stojan generatora jest umieszczony wewnątrz korpusu, który posiada płaszcz chłodzący. Układ łopatkowy mikroturbiny został oddzielony od części generatorowo-łożyskowej za pomocą bezkontaktowego uszczelnienia labiryntowego. Ponieważ w obu częściach korpusu występuje ten sam czynnik roboczy, nie było potrzeby stosowania hermetycznych uszczelnień obrotowych. Niewielkie przecieki wewnątrz korpusu są w tym przypadku dopuszczalne, a nadmiar ciekłego czynnika niskowrzącego jest odprowadzany za pomocą systemu otworów. Zdjęcie kompletnej mikroturbiny zamontowanej na stanowisku badawczym zostało przedstawione na rys. 2. W porównaniu do modelu, przedstawionego na rys. 1, na zdjęciu widać dodatkowe pierścienie połączone trzema prętami gwintowanymi, które stanowią dodatkowe zabezpieczenie. Montaż wirnika w korpusie odbywa się w kierunku osiowym, przy zdjętych pokrywach korpusu oraz zdemontowanej tarczy wirnikowej i tarczy łożyska oporowego. Przed uruchomieniem mikroturbiny w układzie ORC z czynnikiem niskowrzącym została ona przebadana z zastosowaniem sprężonego powietrza jako czynnika roboczego. Pozwoliło to m.in. na sprawdzenie poprawności działania układu wirującego, testy systemu pomiarowego oraz skontrolowanie szczelności korpusu.



Rys. 2. Mikroturbina zainstalowana na stanowisku badawczym

Na poszczególne części mikroturbiny podczas pracy działają różnego typu obciążenia zewnętrzne i wewnętrzne. Ze względu na wysoką prędkość obrotową, na cały układ wirujący (w tym wał, wirnik generatora oraz tarczę wirnikową mikroturbiny) działa siła odśrodkowa. Niewyważenie resztkowe tych elementów powoduje drgania wirnika w kierunku promieniowym i osiowym. Drgania wału mogą być również wzbudzane przez drgania własne tarczy wirnikowej i łopatek. Napływ gorącej pary czynnika roboczego na łopatki mikroturbiny powoduje nagrzewanie się tarczy wirnikowej, wału oraz korpusu. Para jest również dostarczana do łożysk gazowych. Przyczynia się to do powstawania naprężeń cieplnych. Przepływ pary przez układ łopatkowy powoduje również drgania całego wału w kierunku promieniowym i osiowym. Podczas pracy generatora elektrycznego wytwarza się ciepło, które powoduje nagrzewnie się samego generatora oraz wału i korpusu. Ze względu na ograniczoną temperaturę pracy generatora wymagane jest zewnętrzne chłodzenie (płaszczem wodnym), które powoduje wzrost gradientu temperatury i dodatkowo zwiększa

naprężenia cieplne w korpusie. Zmiany temperatury powodują również odkształcenia wirnika i korpusu, co może wpływać na pojawienie się niezgodności geometrycznych i przecieków. Na generator działają także obciążenia elektryczne, związane z występowaniem sił elektromagnetycznych. Wszystkie obciążenia z układu wirującego są przekazywane za pośrednictwem czopów wału na łożyska. Aby łożyska gazowe mikroturbiny poprawnie pracowały, muszą być stale smarowane parą czynnika niskowrzącego. Dzięki temu w łożyskach występuje tarcie płynne, które ogranicza straty tarcia i zapobiega uszkodzeniom powierzchni ślizgowych. Bardzo ważnym zadaniem układu łożysk jest zapewnienie stabilnej pracy wirnika przy minimalnym poziomie drgań. Jest to bardzo istotne przy wysokich prędkościach obrotowych, przy których wzrasta obciążenie wirnika, a w łożyskach mogą pojawić się problemy z niestabilnością, związane z przepływem czynnika smarnego. Wszystkie części znajdujące się wewnątrz korpusu są dodatkowo narażone na działanie czynnika niskowrzącego, który jest bardzo przenikliwy i może mieć destrukcyjny wpływ na niektóre materiały.

Zgodnie z tym co wykazano powyżej, części mikroturbiny muszą poprawnie funkcjonować w bardzo trudnych warunkach. Dlatego podczas eksploatacji tego typu maszyn wymagane jest stosowanie zaawansowanym metod i systemów diagnostycznych, umożliwiających stałą ocenę ich stanu dynamicznego. Jak wykazały wcześniejsze obliczenia symulacyjne, zmiany niektórych parametrów układu wirującego omawianej mikroturbiny mogą mieć bardzo duży wpływ na jej właściwości dynamiczne [30]. Potwierdziły to badania eksperymentalne, które zostały omówione w kolejnej części artykułu.

3. Badania eksperymentalne drgań mikroturbiny

3.1. Badania mikroturbiny bez defektów

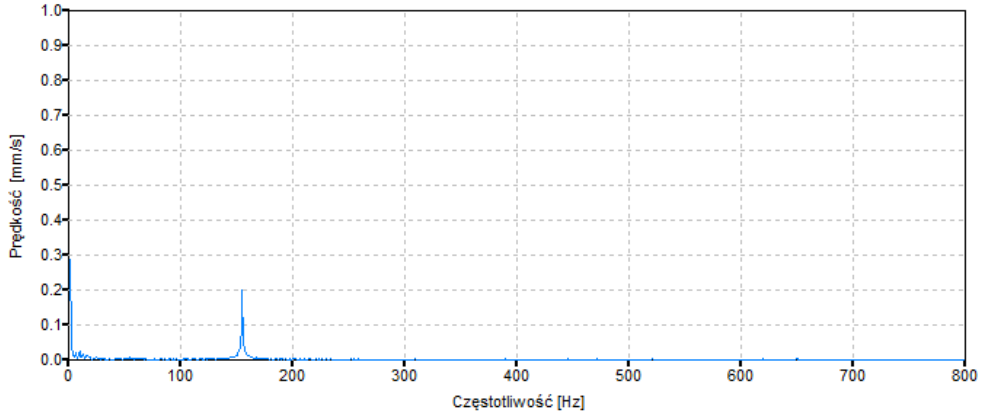
Badania eksperymentalne mikroturbiny zostały wykonane w laboratorium IMP PAN, na stanowisku badawczym umożliwiającym odtworzenie rzeczywistych warunków pracy. Podczas badań układ łopatkowy mikroturbiny oraz łożyska były zasilane parą czynnika niskowrzącego o nazwie HFE-7100, której ciśnienie dochodziło do 11 bar, a temperatura osiągała wartości dochodzące do 180°C. Są to parametry typowe dla docelowych warunków pracy przewidzianych dla małych, domowych układów kogeneracyjnych ORC współpracujących z kotłami wielopaliwowymi. Pomiary wibrodiagnostyczne były prowadzone równolegle z pomiarami parametrów termodynamicznych czynnika roboczego oraz pomiarem prędkości obrotowej wirnika i mocy elektrycznej uzyskiwanej w generatorze. Do pomiaru drgań zastosowano układ pomiarowy składający się z przenośnego analizatora drgań DIAMOND401A w wersji XT (produkowany przez firmę MBJ Electronics) oraz jednoosiowy akcelerometr o oznaczeniu 622B01 firmy PCB Piezotronics. Akcelerometr był mocowany za pomocą podstawki magnetycznej na korpusie mikroturbiny, w obrębie łożyska znajdującego się pomiędzy tarczą wirnikową i generatorem (rys. 1). Ponieważ korpus był wykonany ze stali nierdzewnej, czujniki były mocowane za pomocą dodatkowych podstawek wykonanych ze stali ferromagnetycznej. Aby wybrać odpowiedni punkt pomiarowy sprawdzono poziom i rozkład drgań w różnych miejscach korpusu, ale pomiar w obrębie łożyska przy tarczy wirnikowej w kierunku pionowym charakteryzował się najwyższym poziomem drgań i uznano go za najbardziej miarodajny. Do analizy i wizualizacji wyników pomiarów zastosowano stację roboczą klasy PC z oprogramowaniem MBJLab, dedykowanym do prowadzenia diagnostyki drganiowej różnego typu maszyn, w tym maszyn wirnikowych.

Wyniki pomiarów drgań zostały przedstawione w postaci rozkładów częstotliwościowych prędkości drgań, co znacząco ułatwiało ich interpretację. Ponieważ

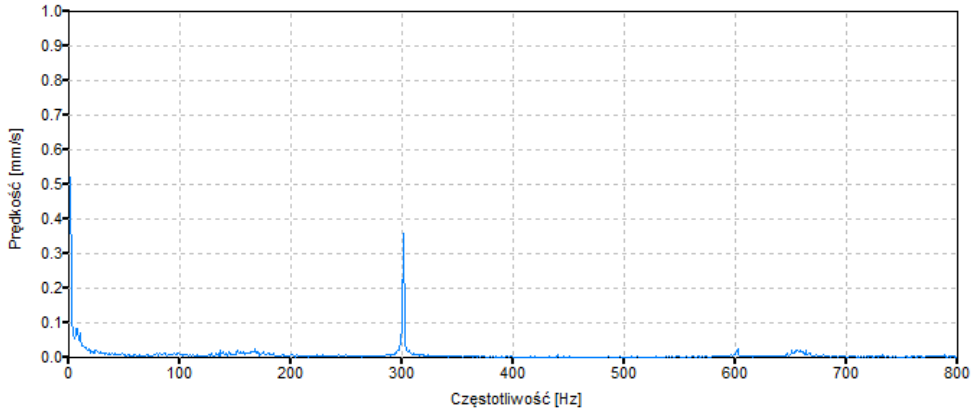
maksymalna prędkość obrotowa badanej maszyny wynosiła aż 24000 obr/min (400 Hz), pomiary były prowadzone w dość szerokim zakresie częstotliwości – od 1 do 800 Hz. Rozdzielczość pomiaru wynikała z maksymalnych możliwości zastosowanego analizatora drgań i wynosiła 1 Hz. Przedstawione w artykule charakterystyki zostały uzyskane jako widma uśrednione z 3 rejestrowanych po sobie widm drgań (zastosowano uśrednianie RMS). Aby ułatwić porównywanie ze sobą wyników pomiarów, dla wszystkich charakterystyk przyjęto wspólną wartość maksymalnej amplitudy prędkości drgań wynoszącą 1 mm/s. Wszystkie rozkłady częstotliwościowe drgań korpusu turbiny zostały uzyskane podczas stabilnej pracy mikroturbiny, tj. przy stałej mocy oraz prędkości obrotowej wirnika. Wykonanie pomiarów przy zmieniających się prędkościach obrotowych wymagałoby zastosowania specjalnych metod analizy sygnałów, dedykowanych do pomiarów w stanach nieustalonych [20]. Utrudniałoby to również bezpośrednie porównanie ze sobą różnych charakterystyk, uzyskanych w zmiennych warunkach.

Wyniki badań przedstawione w tej części artykułu zostały uzyskane podczas jednego z pierwszych testów mikroturbiny, po kilku godzinach pracy od momentu jej uruchomienia. Można więc przyjąć, że była to maszyna nowa, nie posiadająca żadnych defektów. Podczas tych badań prowadzono stały nadzór drganiowy maszyny, ale ze względu na ograniczoną objętość artykułu, w tym miejscu zostały przedstawione tylko wybrane wyniki pomiarów (uzyskane dla różnych prędkości obrotowych), podczas których można było zaobserwować stabilizację prędkości obrotowej, mocy oraz parametrów termodynamicznych pary czynnika roboczego. Rys. 3 przedstawia rozkład częstotliwościowy prędkości drgań uzyskany przy prędkości wirnika równej 9300 obr/min, natomiast na rys. 4 i 5 przedstawiono widma drgań uzyskane dla prędkości 18060 i 21060 obr/min.

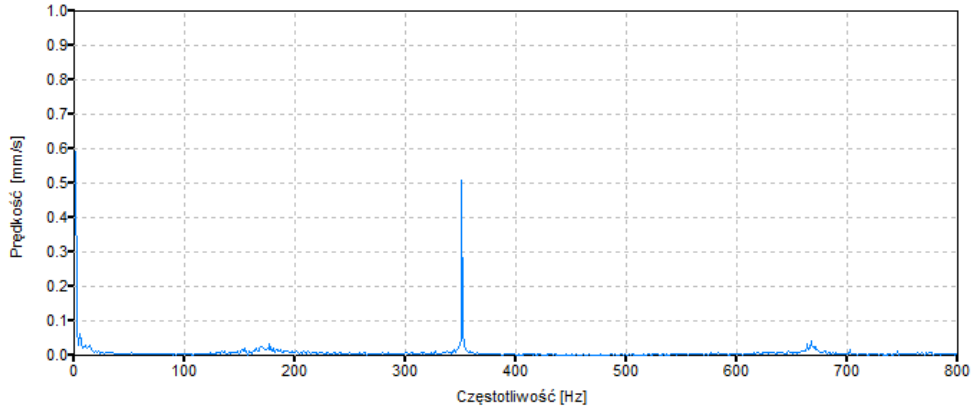
Na wszystkich przedstawionych charakterystykach można zaobserwować jeden dominujący prążek drgań, występujący przy częstotliwości zgodnej z częstotliwością obrotową wirnika (tzw. drgania synchroniczne). Na kolejnych rysunkach były to częstotliwości: 155, 301 i 351 Hz. Amplituda prędkości drgań wraz ze wzrostem prędkości obrotowej wirnika stale rosła, od wartości 0,20 mm/s, poprzez 0,36 mm/s, aż do 0,52 mm/s. Stały wzrost poziomu drgań w analizowanym zakresie prędkości obrotowych jest naturalny, gdyż jest związany ze wzrostem siły odśrodkowej, pochodzącej od niewyważenia resztkowego, wymuszającej drgania wirnika w kierunku poprzecznym. Wirnik badanej mikroturbiny w całym zakresie prędkości roboczych jest wirnikiem podkrytycznym, gdyż uzyskana symulacyjnie prędkość rezonansowa (związana z pierwszą postacią drgań giętnych) występowała dopiero przy ok. 130000 obr/min [30]. Dlatego stopniowy (ale nie nadmierny) wzrost poziomu drgań przy zwiększaniu prędkości obrotowej był spodziewany i nie oznaczał żadnych problemów dynamicznych. W pozostałym paśmie częstotliwości podwyższony poziom drgań można było jeszcze zaobserwować przy najniższych częstotliwościach (do ok. 20 Hz). Nie były to jednak drgania związane z pracą turbiny i występowały one nawet przy zatrzymanym wirniku. W tym zakresie występowały drgania konstrukcji stanowiska badawczego oraz podłoża, na którym było ono ustawione (podłoga laboratorium znajduje się na piętrze, którego strop posiada konstrukcję kratową). Drgania tych elementów były wzbudzane między innymi przez ruch osób znajdujących się w laboratorium i nie udało się ich wyeliminować podczas badań. Dodatkowo, przy prędkościach obrotowych wynoszących 18060 obr/min i 21060 obr/min, można było zaobserwować nieznacznie podwyższony poziom drgań przy częstotliwościach dwa razy większych i dwa razy mniejszych od częstotliwości obrotowej. Są to tzw. składowe 2X oraz 1/2X, ale ze względu na bardzo niskie amplitudy drgań (poniżej 0,05 mm/s) składowe te nie miały istotnego wpływu na stan dynamiczny maszyny.



Rys. 3. Rozkład częstotliwościowy drgań zmierzonych na korpusie mikroturbiny przy prędkości obrotowej wynoszącej 9300 obr/min (155 Hz)



Rys. 4. Rozkład częstotliwościowy drgań zmierzonych na korpusie mikroturbiny przy prędkości obrotowej wynoszącej 18060 obr/min (301 Hz)



Rys. 5. Rozkład częstotliwościowy drgań zmierzonych na korpusie mikroturbiny przy prędkości obrotowej wynoszącej 21060 obr/min (351 Hz)

Oceniając zarejestrowane poziomy drgań w oparciu o normy branżowe, można stwierdzić, że stan dynamiczny badanej mikroturbiny był bardzo dobry. Zgodnie z normą ISO 10816 ogólny poziom wartości skutecznej prędkości drgań (Vrms) mierzony na obudowie łożyska maszyn wirnikowych o mocy do 15 kW nie powinien przekraczać wartości 1,8 mm/s. Ten poziom drgań umożliwia długotrwałą pracę bez żadnych ograniczeń eksploatacyjnych. W omawianym przypadku najwyższa amplituda prędkości drgań wystąpiła przy prędkości 21060 obr/min i osiągnęła zaledwie 0,52 mm/s. Ogólny poziom drgań był niewiele wyższy i mieścił się w zakresie wartości dopuszczalnej dla maszyn oddanych do eksploatacji (0,71 mm/s).

Ponieważ badana mikroturbina pracowała poprawnie w całym zakresie obciążenia i prędkości obrotowej, przedstawione w tej części artykułu wyniki można traktować jako wartości referencyjne. Zostały one uzyskane dla maszyny wirnikowej, której stan techniczny nie budził żadnych zastrzeżeń. Wyniki takich pomiarów mogą być więc wykorzystane jako punkt odniesienia przy dalszych badaniach oraz przy ocenie stanu dynamicznego po pewnym czasie eksploatacji. Uzyskane rozkłady częstotliwościowe drgań mogą być również pomocne przy definiowaniu symptomów różnego typu defektów.

3.2. Badania mikroturbiny z niewyważonym wirnikiem

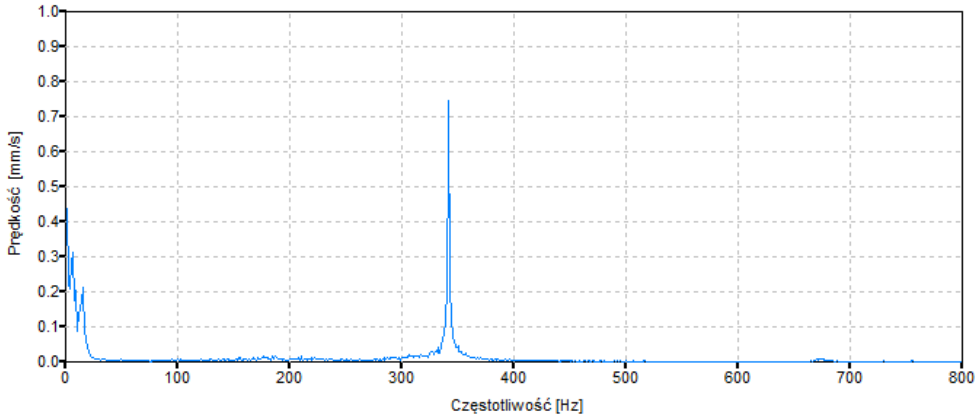
Badana mikroturbina, po badaniach wstępnych, które potwierdziły jej wszystkie parametry projektowe oraz bardzo dobry stan dynamiczny, została poddana kolejnym testom laboratoryjnym. Celem tych badań było przede wszystkim wyznaczenie jej charakterystyk użytkowych (takich jak moc oraz spadki temperatury i ciśnienia czynnika roboczego), ale równolegle podczas tych badań prowadzono również pomiary drgań. Badania wibrodiagnostyczne miały na celu wykrycie ewentualnych problemów dynamicznych mikroturbiny oraz różnego typu defektów, które mogły się pojawić ze względu na doświadczalny charakter układu kogeneracyjnego ORC i testowanie go w różnych, również szybko zmieniających się warunkach [29].

Pierwsze symptomy mogące świadczyć o pogarszaniu się stanu dynamicznego maszyny zostały zaobserwowane już po kilkudziesięciu godzinach pracy. Niepokój wzbudził fakt, że poziom drgań mierzony na korpusie mikroturbiny wzrósł, co było szczególnie widoczne przy wyższych prędkościach obrotowych. Szczegółowa analiza zarejestrowanych rozkładów częstotliwościowych drgań wykazała, że poza składową synchroniczną (1X) w widmie drgań nie pojawiły się dodatkowe składowe, ale sam poziom drgań synchronicznych znacznie wzrósł. Przy prędkości wynoszącej 20520 obr/min amplituda prędkości drgań osiągnęła wartość 0,74 mm/s (rys. 6), a przy prędkości równej 21000 obr/min wynosiła aż 0,87 mm/s (rys. 7). W porównaniu do stanu bazowego oznaczało to wzrost poziomu drgań o około 67%, przy zbliżonej prędkości obrotowej. Poza podwyższonym poziomem drgań nie zaobserwowano żadnych innych objawów, które mogłyby świadczyć o niepoprawnej pracy mikroturbiny.

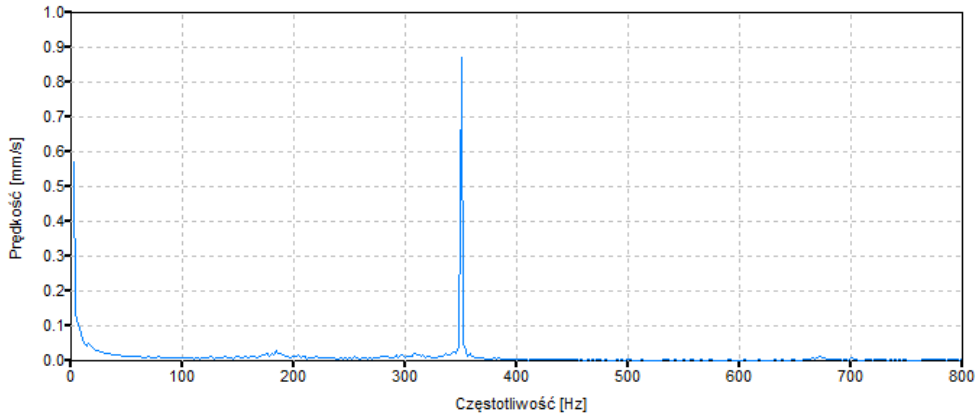
Aby wyjaśnić przyczynę tak dużego wzrostu poziomu drgań podjęto decyzję o demontażu mikroturbiny i przeprowadzeniu wizualnej oceny stanu technicznego jej części. Już na wstępnym etapie demontażu, po odkręceniu pokrywy korpusu od strony nienapędzanej (NDE), można było zaobserwować tarczę oporową łożyska wzdłużnego, która na znacznej powierzchni była pokryta zanieczyszczeniami. Jest to przedstawione na rys. 8. Widoczne na tym samym zdjęciu wgłębienia w tarczy oporowej nie są oznaką uszkodzenia, gdyż zostały wykonane celowo podczas wyważania wirnika. Dalszy demontaż mikroturbiny wykazał, że znaczna część powierzchni wirnika była trwale zanieczyszczona, co powodowało m.in. jego dodatkowe niewyważenie. Zaobserwowany podczas pomiarów drgań wzrost amplitudy składowej synchronicznej (1X) był typowym symptomem zbyt dużego niewyważenia układu wirującego.

Aby znaleźć przyczynę zanieczyszczenia wirnika przeanalizowano przebieg wcześniejszych badań, które były prowadzone na tym samym stanowisku badawczym. Ponieważ jest to stanowisko uniwersalne były na nim wcześniej prowadzone badania różnych podzespołów układu ORC, w tym różnego typu pomp obiegowych. Okazało się, że najbardziej prawdopodobną przyczyną zanieczyszczenia czynnika roboczego układu ORC był olej, który dostał się do obiegu na skutek awarii jednej z badanych pomp. Była to pompa membranowa, której membrany wykonane z tworzywa sztucznego nie wytrzymały długotrwałego kontaktu z czynnikiem niskowrzącym. Na skutek przerwania membrany olej znajdujący się wewnątrz pompy został z niej wypłukany i w całości trafił do obiegu ORC.

Pomimo, że po tym incydencie czynnik niskowrzący układu ORC został wymieniony, to znaczne ilości oleju pozostały wewnątrz rozbudowanej instalacji rurowej oraz w częściach takich jak wymienniki ciepła. Podczas dalszych testów dostał się on razem z czynnikiem niskowrzącym do wnętrza mikroturbiny. Praca mikroturbiny w wysokiej temperaturze spowodowała następnie osadzenie się resztek oleju na jej częściach. Na stan powierzchni wirnika niekorzystny wpływ miał również długotrwały kontakt z mieszaniną czynnika niskowrzącego i oleju. Linia zalania korpusu tą mieszaniną jest wyraźnie widoczna na rys. 8.



Rys. 6. Rozkład częstotliwościowy drgań zmierzonych na korpusie mikroturbiny z niewyważonym wirnikiem przy prędkości obrotowej wynoszącej 20520 obr/min (342 Hz)



Rys. 7. Rozkład częstotliwościowy drgań zmierzonych na korpusie mikroturbiny z niewyważonym wirnikiem przy prędkości obrotowej wynoszącej 21000 obr/min (350 Hz)

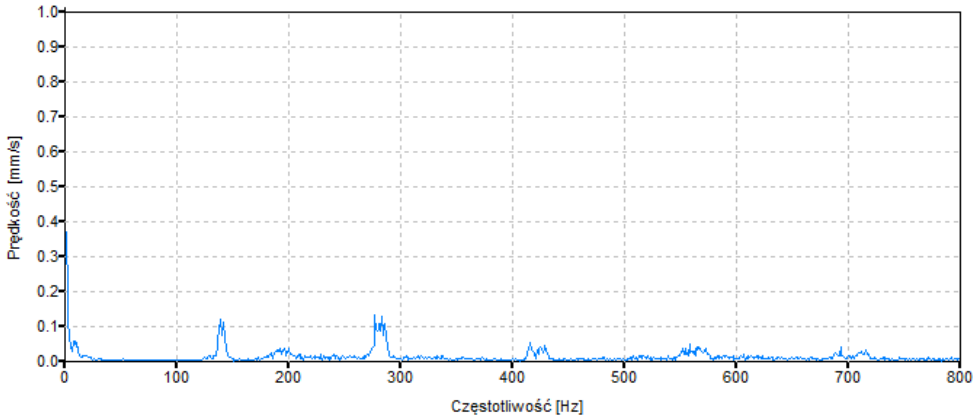
Rys. 8. Tarcza oporowa łożyska z zanieczyszczeniem powodującym niewyważenie wirnika

Po demontażu podzespołów mikroturbiny zostały one poddane czyszczeniu, które umożliwiło pozbycie się wszystkich nalotów oraz zanieczyszczeń. Pozwoliło to na przywrócenie jej dobrego stanu technicznego. Przed kolejnymi badaniami przeprowadzono również płukanie całej instalacji ORC oraz przefiltrowano czynnik niskowrzący. Podjęte działania umożliwiły ponowne uruchomienie stanowiska badawczego z mikroturbiną, a wykonane kontrolnie pomiary drgań potwierdziły jej bardzo dobry stan dynamiczny. Podsumowując tą część artykułu można jeszcze dodać, że pogorszenie się stanu dynamicznego badanej maszyny przepływowej nie było w tym przypadku spowodowane błędami konstrukcyjnymi ani niepoprawnym jej wykonaniem czy montażem, a wynikało wyłącznie z niepoprawnej eksploatacji – zasilania mikroturbiny zanieczyszczonym czynnikiem roboczym.

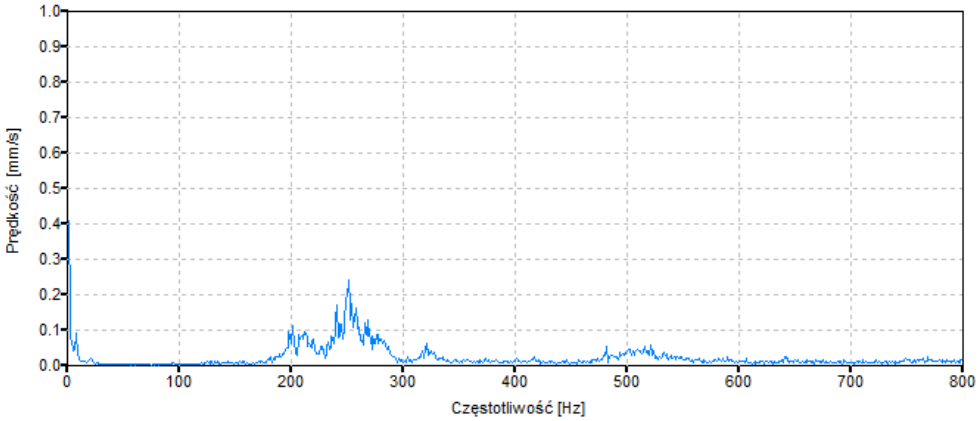
3.3. Badania mikroturbiny z niepoprawnie działającym łożyskiem

Oczyszczenie mikroturbiny i instalacji ORC z pozostałości oleju umożliwiło przywrócenie stanowiska badawczego do stanu pozwalającego na kontynuację badań. Kolejne symptomy niepoprawnej pracy mikroturbiny pojawiły się po jednym z dłuższych (kilkutygodniowych) przestojów. Co prawda, po uruchomieniu mikroturbina pracowała poprawnie, ale tylko przy bardzo niskich prędkościach obrotowych (do ok. 6000 obr/min). Przy wyższych prędkościach w widmie drgań korpusu oprócz składowej 1X pojawiły się wyższe składowe harmoniczne (2X, 3X, 4X oraz kolejne). Dodatkowo, każda z tych składowych nie była skupiona, lecz rozkładała się na kilka prążków o zbliżonych częstotliwościach. Składowa 2X osiągała zbliżoną amplitudę do składowej synchronicznej, a amplitudy kolejnych składowych stopniowo malały. Zostało to przedstawione na rys. 9. Pomimo niepokojącego obrazu drgań kontynuowano badania, rozpędzając wirnik do wyższych prędkości. Przy prędkościach obrotowych powyżej 12000 obr/min wyższe składowe harmoniczne stopniowo zanikały, ale główny prążek drgań (którego częstotliwość była zgodna z prędkością obrotową wirnika) rozpraszał się na szerszy zakres częstotliwości. Widmo drgań stawało się chaotyczne. Zostało to przedstawione na rys. 10. Tego typu widma drgań są typowe dla układów wirujących, w których występuje fizyczny kontakt elementów wirujących z korpusem lub innymi częściami nieobracającymi się. Ponieważ charakterystyka drganiowa maszyny uzyskana w kolejnych kilku próbach nie poprawiała się, podjęto decyzję o konieczności jej demontażu celem dokonania kontroli wszystkich części.

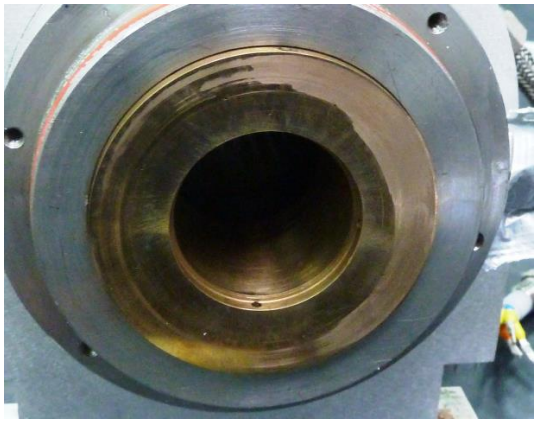
Demontaż mikroturbiny polegał na zdjęciu obydwóch pokryw korpusu (przedniej i tylnej) oraz wyjęciu na zewnątrz wirnika. Oględziny części wykazały, że jedno z gazowych łożysk wzdłużnych nosiło wyraźne ślady zużycia (rys. 11), świadczące o jego niepoprawnej pracy. Było to łożysko znajdujące się na wolnym końcu wału (patrz rys. 1). Niemal na połowie powierzchni ślizgowej łożyska widać było ślady tarcia, które nastąpiło w wyniku fizycznego kontaktu łożyska z powierzchnią tarczy oporowej. Dokładne oględziny łożyska wykazały, że doszło w nim do zatkania 3 z 8 otworów zasilających, przez które pomiędzy powierzchnie ślizgowe dostarczany jest czynnik smarny w postaci pary czynnika niskowrzącego. Są to otwory o średnicy 0,4 mm i małe drobinki różnych osadów, pochodzących z różnych części instalacji doprowadziły do ich zatkania. Niekorzystny wpływ miał tu również dłuższy przestój mikroturbiny, w czasie którego doszło do nagromadzenia się zanieczyszczeń w komorach zasilających łożyska gazowe. W wyniku zatkania otworów zmniejszyła się ilość pary dostarczanej do przestrzeni smarnej, łożyska co doprowadziło do obniżenia nośność. Wraz ze zwiększaniem natężenia przepływu i ciśnienia pary podawanej na układ łopatkowy mikroturbiny rosła również siła osiowa przenoszona przez wał na łożysko oporowe. Niepoprawnie funkcjonujące łożysko nie było w stanie przenieść tej siły, na skutego czego doszło do fizycznego kontaktu dwóch współpracujących ze sobą powierzchni oporowych. W normalnych warunkach pracy są one oddzielone warstwą czynnika smarnego.



Rys. 9. Rozkład częstotliwościowy drgań zmierzonych na korpusie mikroturbiny z uszkodzonym łożyskiem wzdłużnym przy prędkości obrotowej wynoszącej 8340 obr/min (139 Hz)



Rys. 10. Rozkład częstotliwościowy drgań zmierzonych na korpusie mikroturbiny z uszkodzonym łożyskiem wzdłużnym przy prędkości obrotowej wynoszącej 15060 obr/min (251 Hz)

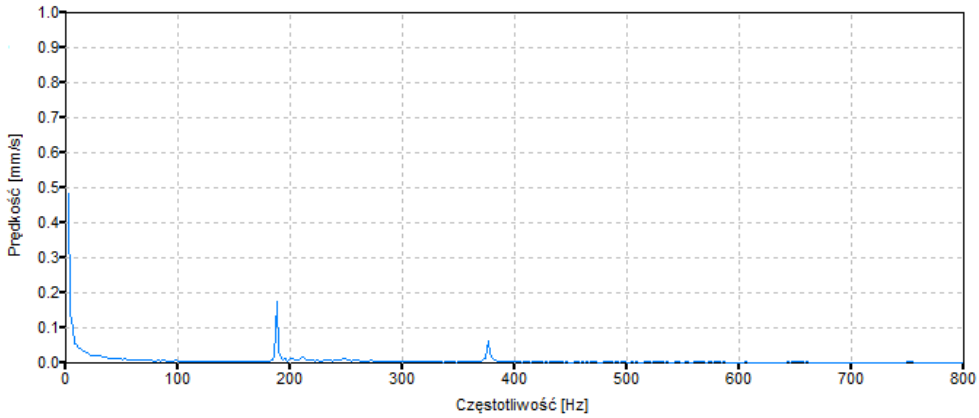


Rys. 11. Gazowe łożysko wzdłużne z uszkodzoną powierzchnią ślizgową

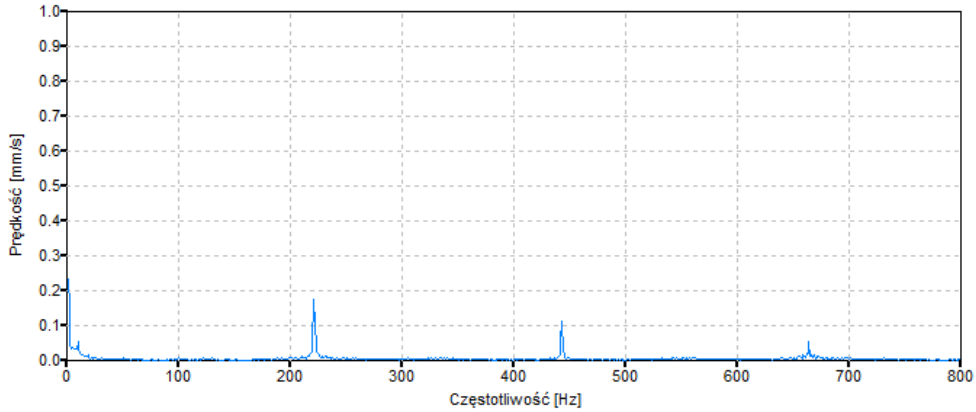
Oględziny wszystkich części mikroturbiny wykazały, że uszkodzeniu uległa tylko powierzchnia ślizgowa jednego łożyska oporowego. Ponieważ tarcza oporowa była wykonana ze znacznie twardszego materiału uszkodzenie wystąpiło tylko na powierzchni panwi łożyska (wykonanej z brązu). Stała kontrola drgań mikroturbiny umożliwiła bardzo szybkie wykrycie tego problemu, dzięki czemu powstałe uszkodzenia mogły być naprawione przez zeszlifowanie z całej powierzchni warstwy materiału o bardzo małej grubości. Po tej naprawie mikroturbina mogła być przywrócona do dalszej eksploatacji, a zmierzony poziom drgań i ich rozkład częstotliwościowy tylko nieznacznie odbiegał od wartości referencyjnych.

3.4. Badania mikroturbiny z wygiętym wałem

Kolejny przypadek niepoprawnej pracy mikroturbiny ORC o mocy 2,5 kW był związany z pojawieniem się w widmie drgań oprócz składowej synchronicznej (1X) dodatkowej składowej harmonicznej o częstotliwości dwa razy większej od aktualnej częstotliwości obrotowej wirnika (2X). Amplituda tej składowej drgań była zazwyczaj o połowę mniejsza niż amplituda składowej 1X i to niezależnie od prędkości obrotowej. Dwa przykładowe widma drgań przedstawiające taki rozkład amplitud drgań zostały przedstawione na rys. 12 i 13. Pomiary drgań zostały wykonane przy prędkościach obrotowych wirnika wynoszących odpowiednio 11280 i 13260 obr/min. W drugim przypadku (rys. 13) oprócz składowej 2X można zaobserwować dodatkowy prążek drgań przy częstotliwości trzy razy większej od częstotliwości podstawowej. Amplituda tej składowej była jednak znaczne niższa od pozostałych i w związku z tym miała ona niewielki wpływ na ogólny poziom drgań, na podstawie którego ocenia się stan dynamiczny maszyn.



Rys. 12. Rozkład częstotliwościowy drgań zmierzonych na korpusie mikroturbiny z wygiętym wałem przy prędkości obrotowej wynoszącej 11280 obr/min (188 Hz)



Rys. 13. Rozkład częstotliwościowy drgań zmierzonych na korpusie mikroturbiny z wygiętym wałem przy prędkości obrotowej wynoszącej 13260 obr/min (221 Hz)

Pojawienie się w widmie drgań dodatkowej składowej 2X, która nie zanika po zmianie prędkości obrotowej, zazwyczaj związane jest ze zgięciem wału lub rozosiowaniem dwóch współpracujących ze sobą wałów. Ponieważ w naszym przypadku w maszynie wirnikowej występuje tylko jeden wał, jako główną przyczynę pojawienia się składowej 2X można było podejrzewać pojawienie się wygięcia wału. Analiza konstrukcji mikroturbiny oraz warunków jej rozruchu i eksploatacji potwierdziła możliwość wystąpienia takiego zjawiska. Wlot pary świeżej do układu łopatkowego mikroturbiny znajduje się po jednej stronie korpusu (patrz

rys. 1), co może powodować nierównomierne nagrzewanie się wirnika. Wirnik tej maszyny zaczyna się obracać dopiero przy odpowiedniej różnicy ciśnień pomiędzy wlotem i wylotem pary. Oznacza to, że w początkowym okresie rozruchu może on być nierównomiernie nagrzewany. Podobna sytuacja występuje w poprzecznych łożyskach gazowych, które są zasilane przez otwory o małej średnicy. Zbyt niskie ciśnienie czynnika smarnego nie spowoduje uniesienia się czopów w łożyskach. Oznacza to, że otwory zasilające znajdujące się w dolnej części łożysk są zasłonięte przez czopy, a przepływ pary odbywa się przez pozostałe otwory. Dopiero po wzroście ciśnienia, powodującym zmianę pozycji czopów w kierunku środka łożysk, przez wszystkie otwory zasilające zaczyna przepływać podobny strumień pary. Sytuację dodatkowo pogarsza fakt, że dwa wloty gorącej pary zasilającej łożyska znajdują się po jednej stronie korpusu. Na etapie rozruchu maszyny może to powodować nierównomierne nagrzewanie się korpusu i związane z tym odkształcenia termiczne. Ponieważ tuleje łożyskowe są osadzone na wcisk w korpusie, odkształcenia korpusu powodują również zmianę geometrii szczelin smarnych w łożyskach. Wszystko to sprawia, że układ wirujący mikroturbiny jest czuły na wszelkiego rodzaju odkształcenia termiczne, które mogą być źródłem symptomów drganiowych typowych dla zgiętego wału.

Szersze badania tego zjawiska wykazały, że dodatkowe składowe drgań (wyraźnie widoczna 2X i znacznie mniejsza 3X) pojawiały się najczęściej przy szybkim rozruchu mikroturbiny i zanikały po kilku minutach pracy. Przy odpowiednio długim wygrzewaniu korpusu, łożysk i wirnika problem ten praktycznie nie występował. W związku z tym, aby uniknąć przejściowych pogorszeń stanu dynamicznego, zmienione zostały ustawienia w układzie regulacji stanowiska badawczego, tak aby rozruch mikroturbiny następował zawsze po czasie zapewniającym równomierny rozkład temperatur.

4. Podsumowanie i wnioski

W artykule zostały omówione badania wibrodiagnostyczne prototypowej mikroturbiny parowej o maksymalnej mocy elektrycznej wynoszącej 2,5 kW. Mikroturbina została opracowana z myślą o zastosowaniu w małych układach kogeneracyjnych ORC, służących do zasilania domów jednorodzinnych w energię cieplną i elektryczną. Badania mikroturbiny wykonano metodą eksperymentalną w warunkach laboratoryjnych. Zastosowane do tego celu stanowisko badawcze pozwoliło na odwzorowanie rzeczywistych warunków pracy oraz zastosowanie docelowego czynnika roboczego. Ponieważ badana maszyna wirnikowa miała dość niską moc i niewielkie wymiary, możliwe było przeprowadzenie badań nawet w obecności różnego typu defektów układu wirującego. Tego typu badania byłyby niemożliwe w przypadku dużych turbin parowych, gdyż w tym przypadku podejrzenie wystąpienia poważniejszego uszkodzenia łożysk czy wirnika wiązałoby się z koniecznością natychmiastowego zaprzestania eksploatacji. Również koszty demontażu dużej turbiny parowej i jej naprawy byłyby niewspółmiernie wyższe. Tymczasem w przypadku małej maszyny przepływowej, nawet pomimo bardzo wysokiej prędkości obrotowej (powyżej 20000 obr/min), możliwe było wykonanie obszernych badań eksperymentalnych, również w przypadkach, gdy jej charakterystyki dynamiczne budziły pewne obawy. Tego typu badania mikroturbin energetycznych nie były dotychczas publikowane w literaturze.

Na podstawie wykonanych badań można stwierdzić, że wszystkie wykryte defekty były związane z niewłaściwą eksploatacją, a nie wadliwą konstrukcją czy też niepoprawnym wykonaniem lub montażem. Wszystkie pojawiające się podczas pracy mikroturbiny niedomagania zostały wykryte przy użyciu diagnostyki drganiowej. W odniesieniu do określonych stanów pracy maszyny wyglądało to następująco:

- Po rozruchu oraz w początkowym okresie eksploatacji mikroturbina charakteryzowała się bardzo niskim poziomem drgań, a w widmie drgań można było zaobserwować

wyłącznie składową związaną z prędkością obrotową wirnika (1X). W odniesieniu do normy ISO 10816 stan dynamiczny badanej maszyny można było ocenić jako bardzo dobry. Drgania korpusu były na poziomie typowym dla maszyn nowych, dopiero oddanych do eksploatacji.

- Zanieczyszczenie czynnika roboczego układu ORC olejem (w wyniku awarii pompy) spowodowało przedostanie się oleju do wnętrza korpusu mikroturbiny. Było to przyczyną powstania osadów na częściach wirnika, co spowodowało m.in. znaczne niewyważenie wirnika. Podczas pomiarów diagnostycznych objawiało się to wyraźnym wzrostem poziomu drgań synchronicznych (o ponad 60%), których częstotliwość była zgodna z częstotliwością obrotową wirnika (1X).
- W przypadku niepoprawianie działającego łożyska wzdłużnego, w którym doszło do zatkania kilku otworów zasilających, powyżej pewnych prędkości obrotowych zarejestrowane zostało chaotyczne widmo drgań, które mogło świadczyć o otarciach pomiędzy układem wirującym oraz częściami nieobracającymi się. Początkowo (powyżej 6000 obr/min) pojawiły się wyższe składowe harmoniczne o dość nieregularnym kształcie, ale powyżej 12000 obr/min obserwowany był nietypowy, chaotyczny obraz drgań w obrębie składowej synchronicznej.
- Kolejnym wykrytym defektem mikroturbiny było zgięcie wirnika. Problem ten występował jednak tylko w pierwszych chwilach po rozruchu. Przeprowadzona analiza wykazała, że był on prawdopodobnie związany z nierównomiernym nagrzewaniem się wirnika i korpusu. Zjawisku temu można zapobiec poprzez odpowiednie zaplanowanie procedury wygrzewania maszyny przed rozruchem.

Ponieważ wszystkie defekty, które pojawiły się podczas eksploatacji mikroturbiny, zostały bardzo szybko wykryte, nie doszło do poważniejszych uszkodzeń maszyny. Dlatego w dość krótkim czasie oraz przy niewielkim nakładzie finansowym mikroturbina mogła być przywrócona do pracy na stanowisku badawczym, bez żadnych ograniczeń eksploatacyjnych. Należy również zwrócić uwagę na fakt, że tylko w przypadku niewyważenia wystąpił wyraźny zrost poziomu drgań. Pozostałe defekty zostały wykryte dzięki analizie uzyskiwanych widm drgań. Ogólny poziom drgań w przypadku niesprawności łożyska gazowego oraz wygięcia wirnika pozostawał na niskim poziomie. Proste metody diagnostyczne, oparte wyłącznie o bieżącą kontrolę poziomu drgań, byłyby w tym przypadku zawodne i mogłoby dojść do poważnego uszkodzenia mikroturbiny.

Wykryte podczas badań laboratoryjnych defekty oraz ich symptomy diagnostyczne zostaną wykorzystane przy tworzeniu sytemu diagnostycznego, który będzie niezbędny w docelowym miejscu pracy mikroturbiny. Z wyników badań eksperymentalnych przedstawionych w niniejszym artykule mogą również skorzystać inni inżynierowie i badacze, zajmujący się eksploatacją i diagnostyką różnego typu mikroturbin energetycznych, w tym wysokoobrotowych mikroturbin ORC, stosowanych w coraz bardziej popularnych małych układach kogeneracyjnych.

Podziękowania: Omawiane badania zostały wykonane w ramach projektu nr POIG.01.01.02-00-016/08 oraz w ramach działalności statutowej IMP PAN w Gdańsku. Autorzy artykułu chcieliby podziękować kierownikowi ww. projektu prof. Janowi Kicińskiemu, za inspirację do zajęcia się badaniami układów kogeneracyjnych i mikroturbin energetycznych. Wyrazy podziękowania kierujemy również do prof. Zbigniewa Kozaneckiego, który kierował pracami konstrukcyjnymi nad badaną mikroturbiną, oraz dr. Eugeniusza Ihnatowicza, za jego wkład w budowę Laboratorium Mikrosiłowni Kogeneracyjnych w IMP PAN w Gdańsku.

Literatura

1. Barella S, Bellogini M, Boniardi S, Cincera S. Failure analysis of a steam turbine rotor. Engineering Failure Analysis 2011; 18: 1511-1519.

2. Barsali S, De Marco A, Giglioli R, Ludovici G, Possenti A. Dynamic modelling of biomass power plant using micro gas turbine. Renewable Energy 2015; 80: 806-818.

3. Beith R. (ed.), Small and micro combined heat and power (CHP) systems. Cambridge: Woodhed Publishing Limited, 2011.

4. Czmochowski J, Moczko P, Odyjas P, Pietrusiak D. Tests of rotary machines vibrations in steady and unsteady states on the basis of large diameter centrifugal fans. Eksploatacja i Niezawodność – Maintenance and Reliability 2014; 16(2): 211-216.

5. Dominiczak K, Rządkowski R, Radulski W, Szczepanik R. Online prediction of temperature and stress in steam turbine components using neural network. Journal of Engineering for Gas Turbines and Power 2016; 138: 052606-1.

6. Drescher U, Bruggemann D. Fluid selection for the Organic Rankine Cycle (ORC) in biomass power and heat plants. Applied Thermal Engineering 2007; 27(1): 223-228.

7. Efimov N, Papin V, Bezuglov R. Determination of rotor surfacing time for the vertical microturbine with axial gas-dynamic bearings. Procedia Engineering 2016; 150: 294-299.

8. Hong D, Joo D, Woo B, Koo D, Ahn C. Unbalance Response Analysis and Experimental Validation of an Ultra High Speed Motor-Generator for Microturbine Generators Considering Balancing. Sensors 2014; 14: 16117-16127.

9. Kaczmarczyk T, Żywica G, Ihnatowicz E. Vibroacoustic diagnostics of a radial microturbine and a scroll expander operating in the organic Rankine cycle installation. Journal of Vibration Engineering 2016; 18(6): 4130-4147.

10. Kataoka T, Kishikawa T, Sakata S, Nakagawa T, Ishiguro J. Remote monitoring and failure diagnosis for a microturbine cogeneration system. ASME Turbo Expo 2007, Montreal (Canada), GT2007-27355.

11. Keshtkar H, Alimardani A, Abdi B. Optimization of rotor speed variations in microturbines. Energy Procedia 2011; 12: 789-798.

12. Kiciński J, Żywica G. Steam microturbines in distributed cogeneration, Cham: Springer, 2014.

13. Klonowicz P, Witanowski Ł, Jędrzejewski Ł. A turbine based domestic micro ORC system. Energy Procedia 2017; 129: 923-930.

14. Kozanecka D, Kozanecki Z, Tkacz E, Łagodziński J. Experimental research of oil-free support systems to predict the high-speed rotor bearing dynamics. International Journal of Dynamics and Control 2015; 3(1): 9-16.

15. Kozanecki Z, Łagodziński J. Magnetic thrust bearing for the ORC high - speed microturbine. Solid State Phenomena 2013; 198: 348-353.

16. Kubitz L, Rządkowski R, Gnesin V, Kolodyazhnaya L. Direct integration method in aeroelastic analysis of compressor and turbine rotor blades. Journal of Vibration Engineering & Technologies 2016; 4(1): 37-42.

17. Liu C, Jiang D, Chen J, Chen J. Torsional vibration and fatigue evaluation in repairing the worn shafting of the steam turbine. Engineering Failure Analysis 2012; 26: 1-11.

18. Margo P, Luck R. Energetic and exergetic analysis of waste heat recovery from a microturbine using organic Rankine cycles. International Journal of Energy Research 2013; 37(8): 888-898.

19. Otsu Y, Somaya K, Yoshimoto S. High-speed stability of a rigid rotor supported by aerostatic journal bearings with compound restrictors. Tribology International 2011; 44: 9-17.

20. Pawlik P. Single-number statistical parameters in the assessment of the technical condition of machines operating under variable load. Eksploatacja i Niezawodność – Maintenance and Reliability 2019; 21(1): 164-169.

21. Peirs J, Reynaerts D, Verplaetsen F. A microturbine for electric power generation. Sensors and Actuators 2004; 113: 86-93.

22. Poursaeidi E, Mohammadi Arhani M. Failure investigation of an auxiliary steam turbine. Engineering Failure Analysis 2010; 17: 1328-1336.

23. Poursaeidi E, Taheri M, Farhangi A. Non-uniform temperature distribution of turbine casing and its effect on turbine casing distortion. Applied Thermal Engineering 2014, 71: 433-444.

24. Wang W, Buhl P, Klenk A, Liu Y, The effect of in-service steam temperature transients on the damage behavior of a steam turbine rotor. International Journal of Fatigue 2016; 87: 471-483.

25. Witek L, Orkisz M, Wygonik P, Musili D, Kowalski T. Fracture analysis of a turbine casing. Engineering Failure Analysis 2011; 18: 914-923.

26. Zhang D, Xie Y, Feng Z. An investigation on dynamic characteristics of a high speed rotor with complex structure for microturbine test rig. ASME Turbo Expo 2008, Berlin (Germany), GT2008-50411.

27. Ziegler D, Puccinelli M, Bergallo M, Picasso A. Investigation of turbine blade failure in a thermal power plant. Case Studies in Engineering Failure Analysis 2013; 1: 192-199.

28. Żywica G, Bagiński P. Investigation of gas foil bearings with an adaptive and non-linear structure. Acta Mechanica et Automatica 2019; 13(1): 5-10.

29. Żywica G, Kaczmarczyk T, Ihnatowicz E, Turzyński T. Experimental investigation of the domestic CHP ORC system in transient operating conditions. Energy Procedia 2017; 129: 637-643.

30. Żywica G, Kiciński J. The influence of selected design and operating parameters on the dynamics of the steam micro-turbine. Open Engineering 2015; 5: 385-398.

Aleksander Waliszyn, Andrzej Adamkiewicz
Maritime University of Szczecin, Faculty of Marine Engineering, Wały Chrobrego 1-2, 70-500 Szczecin, Poland
a.valishin@am.szczecin.pl, a.adamkiewicz@am.szczecin.pl

BADANIA ODPORNOŚCI EROZYJNEJ NIKLU I JEGO STOPU DO ZASTOSOWANIA W ELEMENTACH MASZYN PRZEPŁYWOWYCH

STUDIES ON RESISTANCE TO EROSION OF NICKEL AND ITS ALLOYS TO BE USED IN ELEMENTS OF FLUID - FLOW MACHINES

Streszczenie. W artykule przedstawiono wyniki badań odporności metali na uszkodzenia erozyjne zachodzące pod wpływem kawitacji hydraulicznej. Na podstawie wyników wcześniejszych badań, przyjęto hipotezę o zmęczeniowym charakterze zużycia erozyjnego oraz zależności odporności metali na zniszczenia erozyjne od struktury ich sieci krystalicznej. Dla potwierdzenia przyjętej hipotezy na stanowisku kawitacyjno-udarowym sprawdzono metale z różnymi sieciami krystalicznymi: stal 45 (sieć płasko centralna), nikiel 200/201 oraz stop niklu Monel 400 (sieć heksagonalna). Otrzymane wyniki badań potwierdziły przyjętą hipotezę, wskazując tym samym na zasadność stosowania niklowych powłok ochronnych w maszynach przepływowych.

Słowa kluczowe: zniszczenia erozyjne, odporność, stopy niklu, maszyna przepływowa.

1. Fizyka erozji kawitacyjnej w układach chłodzenia silników o zapłonie samoczynnym

Uszkodzenia korozyjno-erozyjne powierzchni maszyn przepływowych, omywanych płynami i powierzchni wymiany ciepła chłodzonych cieczami, w znacznym stopniu pogarszają stan techniczny powierzchni i zmniejszają ich trwałość. Przyczyną powstawania erozji najczęściej jest kawitacja płynu w przestrzeni roboczej maszyny. Pomimo dużej liczby badań nad zjawiskiem kawitacji i powstawaniem uszkodzeń erozyjnych powierzchni metali, otwartą pozostaje fizyka oddziaływania strumienia cieczy na powierzchnie ochronne warstw, a w szczególności warstw tłumiących [3, 4, 7].

Współcześnie zdobyła uznanie teoria korozyjno-erozyjnego niszczenia tulei cylindrowych silników o zapłonie samoczynnym [1, 2]. Przeważającym jest pogląd, że praprzyczyną zniszczeń powierzchni chłodzonych tulei i bloków cylindrowych jest

gwałtowne oddziaływanie cieczy na warstwę wierzchnią metalu, w wyniku implozji pęcherzyków kawitacyjnych [1, 13]. Erozyjne zniszczenie tulei i bloków cylindrowych, przejawiające się powstaniem skupisk głębokich wżerów, odbywa się wskutek złożonego współdziałania procesów mechanicznego i elektrochemicznego uszkadzania metalu elementu, a mianowicie erozji kawitacyjnej i korozji elektrochemicznej. W wyniku oddziaływania uderzeń tłoka, przy przejściach korbowodu przez górny (GMP) i dolny (DMP) martwe punkty, tuleja cylindrowa podlega drganiom o wysokiej częstotliwości, co powoduje zmiany prędkości strumieni cieczy chłodzącej na powierzchniach tulei i bloków cylindrowych. Powstające przy tym lokalne depresje i wzrosty ciśnienia cieczy sprzyjają zrywaniu ciągłości strumienia i tworzeniu w obszarach obniżonego ciśnienia dużej liczby pęcherzyków kawitacyjnych wypełnionych parą, gazem lub ich dwufazową mieszaniną.
Zmianę ciśnienia w dowolnym punkcie na powierzchni tulei omywanej przez ciecz można ocenić bezwymiarowym współczynnikiem miejscowego rozładowania [2]

$$\xi = \left(\frac{\vartheta_0}{\vartheta_i}\right)^2 - 1 \tag{1}$$

gdzie: ϑ_0 – średnia prędkość cieczy omywającej powierzchnię tulei;
ϑ_i – prędkość strumienia w punkcie „i” na powierzchni tulei.
Największe rozładowanie p_i wystąpi w punkcie, w którym współczynnik ξ będzie maksymalny:

$$p_i = p_0 - q\xi \tag{2}$$

gdzie: p_0 – ciśnienie średnie cieczy;
q – prędkość wzrostu ciśnienia strumienia cieczy,

$$q = p\vartheta_i^2/2 \tag{3}$$

ρ – gęstość cieczy.
Proces powstania kawitacji zaczyna się, kiedy p_i osiągnie wartość ciśnienia pary nasyconej cieczy p_n, przy temperaturze otoczenia. Tworzenie pęcherzyków kawitacyjnych następuje w strumieniu opływającym powierzchnie tulei cylindrowych w obszarze, gdzie strumień cieczy przepływa przez przewężenia i ma największą prędkość, a ich implozja zachodzi w zakresie prędkości obniżonej gdzie strumień napotyka opór ruchu lub przepływa przez przestrzenie o rozszerzającym się przekroju. Prędkość przepływu wody w układach chłodzenia okrętowych silników spalinowych na ogół nie przekracza 2 m/s i nie tworzy warunków do powstawania kawitacji hydrodynamicznej. Dlatego, za podstawową przyczyną zrywania ciągłości strumieni cieczy należy uznać drgania tulei cylindrowych o wysokiej częstotliwości [14]. Obecność drgań wysokiej częstotliwości tulei sprzyja powstaniu warunków kawitacji w układach wody chłodzącej. Wyjaśnienie teoretyczne tego zjawiska uzasadnia fakt, że przy zwiększeniu ciśnienia dynamicznego ($q = \rho\vartheta_i^2/2$) maleje cieśnienie statyczne cieczy i tworzą się warunki sprzyjające zerwaniu jej ciągłości. Do tego przyczynia się oddziaływanie rozciągające tulei cylindrowej, na skutek ruchu drgającego.
Ciecz posiada wytrzymałość objętościową i przeciwstawia się naprężeniom rozciągającym. W momencie równości lub przekraczania wytrzymałości objętościowej

pod wpływem naprężeń rozciągających tulei cylindrowej, zaczynają się procesy kawitacyjne – tworzenie pęcherzyków parowo-gazowych. Częstotliwość pulsacji pęcherzyków jest równa wówczas częstotliwości drgań tulei cylindrowej.
Podczas implozji pęcherzyka parowo-gazowego, na granicy jego powierzchni powstają strumienie cieczy o dużej prędkości do 34 m/s, przy czym ciśnienie cieczy na granicy implozji pęcherzyka z powierzchnią metalu omywanego wodą może przekraczać ($5 \cdot 10^6$) Pa [12]. Wówczas energia uderzenia cząstek strumienia kumulatywnego jest równa:

$$E = fm\vartheta_k^2 \left[\frac{\eta_1 \omega}{6kT} + \frac{\eta_2}{c^2 \rho}\right]^{-1}, \tag{4}$$

gdzie: m – masa cieczy uderzającego strumienia,

ϑ – prędkość strumienia kumulatywnego,

η_1 i η_2 – współczynniki lepkości kinematycznej cieczy do i po implozji pęcherzyka,

ω – współczynnik tarcia wewnętrznego molekuł cieczy,

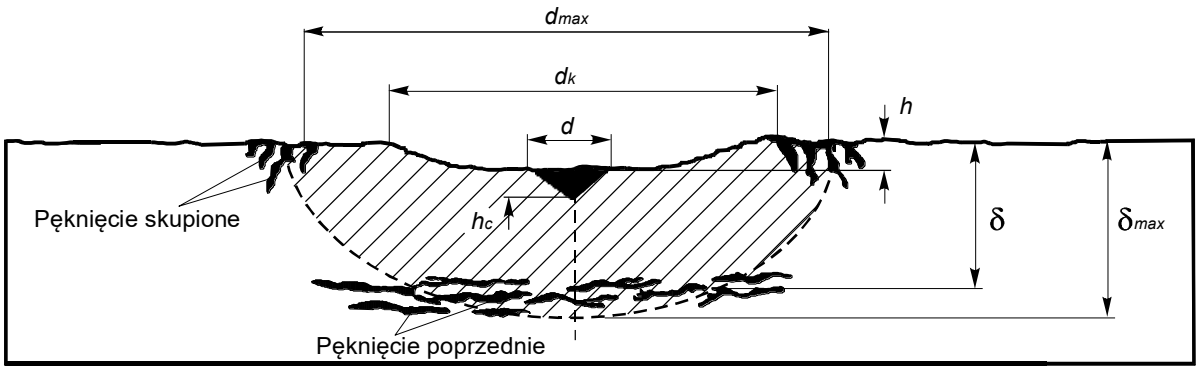
k – stała Boltzmanna,

T – temperatura,

c – prędkość bezwzględna dźwięku,

f – współczynnik uwzględniający okoliczności do uderzenia pęcherzyka.

Oddziaływanie uderzeniowe skumulowanego strumienia cieczy na powierzchnię metalu części opływanej, prowadzi do jej odkształcenia plastycznego i zwiększenia jej twardości – zgniotu [7]. We wzmocnionych warstwach wierzchnich metalu poddawanych dalszemu oddziaływaniu uderzeniowemu skumulowanych strumieni cieczy w rezultacie zmęczenia powstają mikropęknięcia, których dalszy rozwój prowadzi do pojawienia się wżerów erozyjnych, w postaci kraterów. Na rysunku 1 przedstawiono schemat uszkadzania metalu wskutek skumulowanego oddziaływania strumienia [2].



Rys. 1. Schemat uszkadzania metalu w wyniku odziaływania erozji kawitacyjnej
d_{max} – średnica maksymalna, odpowiadająca obszarowi powstawania pęknięć skupionych; d_k – średnica wżeru; d – średnica obszaru powstania produktów utlenienia; δ -głębokość niszczenia metalu

Uszkodzenie metalu następuje po osiągnięciu wartości krytycznych naprężeń, co jest charakterystycznym dla zniszczeń zmęczeniowych materiałów. Często dla zapobieżenia uszkodzeniom erozyjnym w literaturze technicznej proponowane są mało skuteczne metody podwyższenia twardości powierzchni podlegających erozji kawitacyjnej. Jednak w pracy [12] intensywność erozji kawitacyjnej określono zależnością:

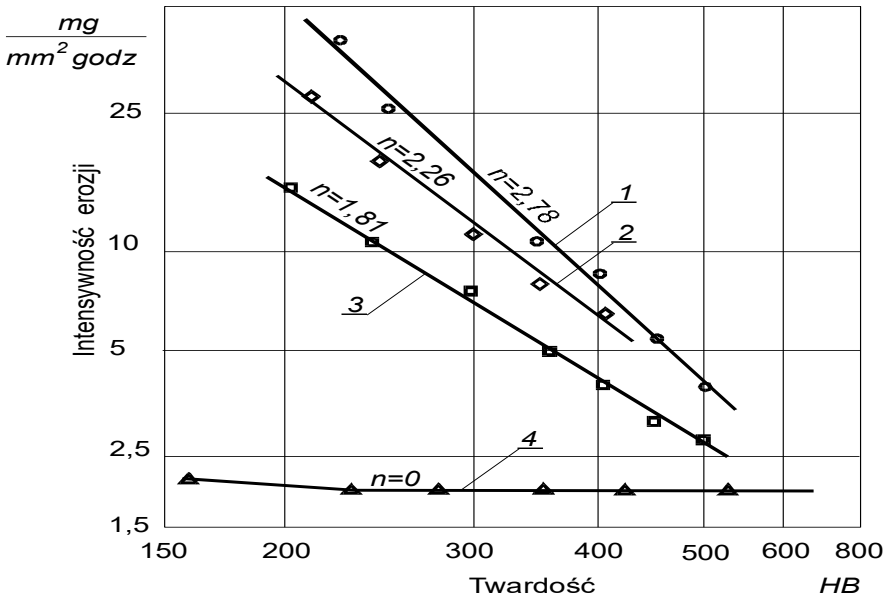
$$J = const \cdot H^n \tag{5}$$

gdzie: J – intensywność zużycia erozyjnego (mg/mm^2 godz) metalu, zdetrminowana nie tylko jego odpornością na niszczenie erozyjne;
H – twardość powierzchni metalu (HB);
n – wykładnik o wartościach od 2,78 (dla stali węglowych) do 0 (dla stali stopowych chromu).

Wpływ początkowej twardości stali na intensywność erozji kawitacyjnej przedstawiono na rys. 2 [12, 13].

Erozyjne zniszczenia stali stopowej 03HG10-10 przy mniejszej twardości w porównaniu ze stalami wykorzystanymi w eksperymencie było znacznie mniejsze. Wynika z tego, iż dla zwiększenia odporności na erozję kawitacyjną powierzchnia metali musi być bardziej plastyczna. Zwiększenie odporności na erozję kawitacyjną powierzchni metali można osiągnąć poprzez:

– laminaryzację strumieni cieczy roboczej;
– pokrycie powierzchni metali powłokami ochronnymi;
– tłumienie drgań części/detalu przy kawitacji wibracyjnej.



Rys. 2. Intensywność erozji kawitacyjnej w zależności od początkowej twardości stali: 1– S40 i S1Cr40; Cr5V3; 2 – 1Cr13, 2Cr13, Mn20Si1; 3 – 1Cr18Ni3Mn4Cu2; 4 – 30Cr10Mn10

Laminarność strumieni cieczy jest związana z powstawaniem układów hydraulicznych, w których prędkości strumieni cieczy nie przekraczają wartości krytycznych,

odpowiadających granicznym wartościom kryterium Reynoldsa. Wartość krytyczną, ustalającą początek kawitacji określa związek matematyczny w postaci [2]:

$$K_k = \frac{K_k^0}{C_j^0}\,(C_j - \Delta C_j^*) \tag{6}$$

gdzie: K_k^0 i C_j^0 – liczba kawitacji i współczynnik oporu nie laminarnej warstwy przyściennej strumienia cieczy;

ΔC_j^* – zależnie od struktury strumienia cieczy określa ją zależność:

$$\Delta C_j^* = \frac{1{,}33\left(1-\frac{1}{e^F}\right)}{\sqrt{Re}} \tag{7}$$

gdzie: e^F - plastyczność graniczna powierzchni omywanego metalu cieczą [2].

Analiza wyrażeń (6-7) dla współczynnika oporu C_j^* wskazuje, że przy zmniejszeniu liczby kryterialnej Reynoldsa i zwiększaniu plastyczności omywanej powierzchni, opór warstwy wierzchniej wzrasta. Odpowiednio obniża się liczba kawitacji K_k, charakteryzująca warunki początku zjawisk kawitacyjnych i związanych z nimi uszkodzeń.

Biorąc pod uwagę że w początkowym momencie, w wyniku oddziaływania strumieni kumulacyjnych w warstwie wierzchniej metali powstaje zgniot, który w warunkach cyklicznych uderzeń niszczy się i pęka. Powstające mikropęknięcie są zarodkami głębokich wżerów przy powstającej wówczas korozji elektrochemicznej szczelinowej. W konsekwencji zniszczenie erozyjno-korozyjne można traktować jako proces zmęczeniowo-korozyjny. Tym samym można stwierdzić, że odporność metali na zniszczenie erozyjno-korozyjne w znacznym stopniu zależy od własności metalu, składu chemicznego i własności otaczającego środowiska.

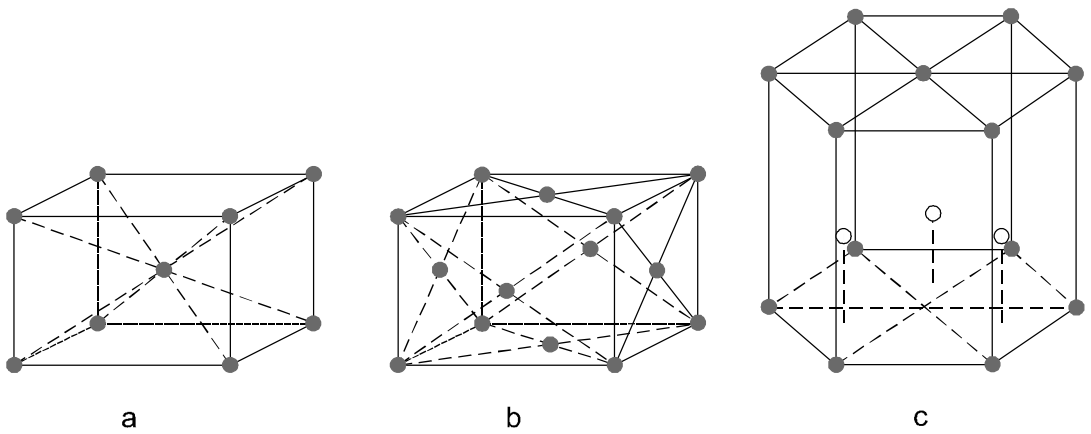
Biorąc pod uwagę zmęczeniowy charakter erozyjnego niszczenia metali, który powstaje w wyniku odziaływań uderzających strumieni cieczy w momencie eksplozji kawitacyjnych pęcherzów parowo-gazowych, można wyciągnąć następujące wnioski. Utwardzanie warstwy powierzchniowej metalu i wynikający z tego zgniot, jak widać z rys. 1, jest początkowym stadium zniszczenia erozyjnego. Ponieważ w dalszym rozwoju procesu, w wyniku powstających naprężeń zmęczeniowych, utwardzona powierzchnia metalu ulega pękaniu, a następnie, powstaniu wżerów erozyjnych, gdy jednocześnie występują naprężenia zmęczeniowe i korozja szczelinowa. Wychodząc z tego, najskuteczniejszej metody zapobiegającej niszczeniu erozyjnemu powierzchni metalowych należy poszukiwać w racjonalnym doborze właściwości mechanicznych metali stosowanych w instalacjach hydraulicznych. Oczywiście głównym warunkiem zastosowania takich metali powinna być zdolność ich do nie tworzenia utwardzonych warstw powierzchniowych, to znaczy powinna być plastyczna.

2. Wybór obiektu badań

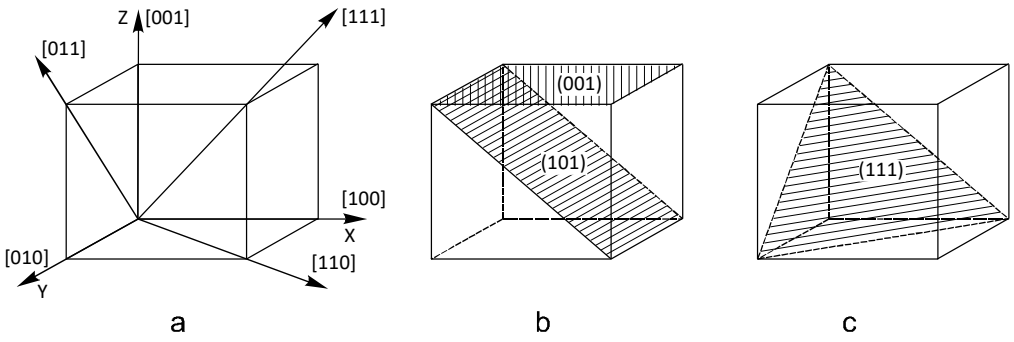
Rozważając zachowanie sprężystości pojedynczych kryształów metali konstrukcyjnych, należy wziąć pod uwagę anizotropię sprężystych modułów krystalicznych, pod kątem określenia różnic w odkształceniu sprężystym, gdy obciążenia są przykładane w różnych kierunkach krystalograficznych. Większość

metali krystalizuje się w trzech typach sieci: sieć przestrzennie centryczna, sieć płasko centryczna i sieć heksagonalna. Takie rodzaje siatek pokazano na rys.3. We wszystkich sieciach krystalicznych indeksy ścian kryształów oznaczono następująco: wzdłuż osi X jako [100]; osi Y - [010] i osi Z - [001]. Indeksy przekątnych ścian oznaczono jako: [110] na płaszczyźnie X-Y , [011] na płaszczyźnie Y-Z i [101] na płaszczyźnie X-Z. Przekątną przestrzenną oznaczono jako [111]. Do metali posiadających strukturę sieci krystalicznej, odpowiadającej sieci płasko centrycznej można zaliczyć żeliwo, stal i miedź, a do metali ze siecią heksagonalną – nikiel, cynk, aluminium i kadm. Ważną charakterystyką metali z punktu widzenia odporności erozyjnej jest ich plastyczność, na co wskazuje wyrażenie (7). W pracy [13] zauważono, że plastyczność metali z siecią heksagonalną jest większa niż metali z siecią płasko centryczną. Odkształcenie tych metali powstaje w postaci poślizgu wzdłuż płaszczyzny [001], co nie powoduje utwardzenia warstwy wierzchniej, a mianowicie: zgniotu. Odkształcenie metali z siecią płasko centryczną w płaszczyznach [111] i [112] przedstawia rysunek 4.

Porównanie tych wartości sugeruje, że nikiel i aluminium są bardziej plastyczne niż stal. Przykłady komórek gęsto upakowanych sieci krystalicznych ścian kryształów przedstawiono na rys. 3, a indeksy kierunków odkształceń komórek krystalograficznych i płaszczyzn ścinania pokazano na rys.4.



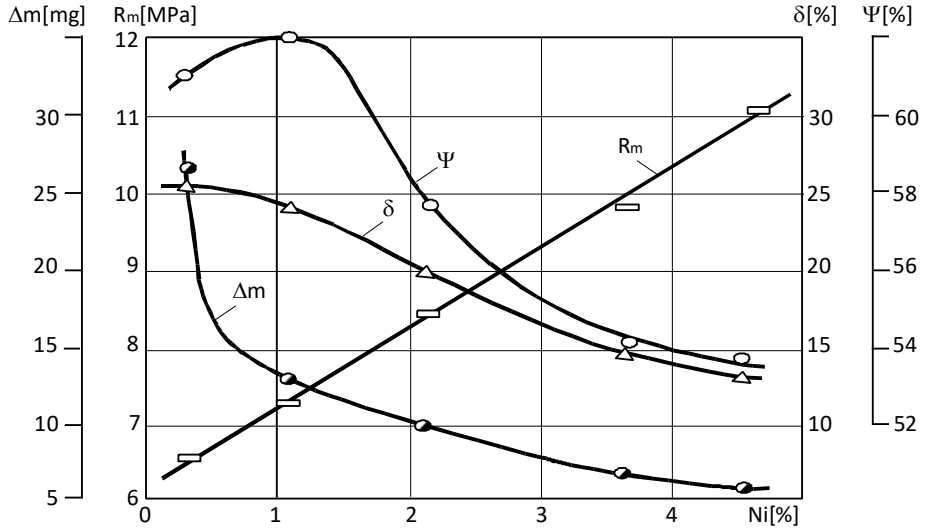
Rys. 3. Elementarne komórki gęsto upakowanych sieci krystalicznych:
a – sieć przestrzennie centryczna; b – sieć płasko centryczna; c – sieć heksagonalna



Rys. 4. Krystalograficzne indeksy kierunków odkształceń (a)
i płaszczyzn ścinania (b, c)

Rozpatrując właściwość plastyczności metali, wzięto pod uwagę wpływ kształtu ich elementarnych komórek sieci krystalicznych i rozmiar ich ziaren. W szczególności, wartość plastyczności $\sigma_{0,2}$ jest różna dla metali, w zależności od rodzaju ich sieci krystalicznej. W przypadku metali o objętościowej siatce krystalicznej (Fe, Cr, Mo, W), a zwłaszcza opartych na ich stopach, w normalnych warunkach występują tylko ułamki procenta. W przypadku metali z siecią płasko centryczną (Cu, Ni, Al, Ag), plastyczność ich może być ponad przecięna. W związku z tym, interesującym jest, zastosowanie w charakterze powłoki ochronnej przed uszkodzeniami erozyjnymi, metali o płasko centrycznej siatce krystalicznej. Oznacza to zastosowanie metali, w których nie powstają znacznie utwardzania warstwy powierzchniowej pod wpływem oddziaływań mikro uderzeniowych.

Wyróżniającym się z tego punktu widzenia materiałem jest nikiel i jego stopy. Nikiel jest bardzo plastycznym metalem, o dobrych własnościach mechanicznych. Stopy niklu: takie jak brązy i mosiądze niklowe, a także związki międzymetaliczne na bazie niklu i aluminium (Ni_3Al i NiAl) mogą być skutecznymi materiałami na powłoki ochronne powierzchni narażonych na kawitację w układach hydraulicznych maszyn przepływowych. Skuteczność chemicznego nakładania powłok niklowych na próbki żeliwa podano w pracach [1,16], gdzie w instalacji hydraulicznej stanowiska badawczego wyposażonego w wibrator magnetostrykcyjny, powierzchne próbek nie wykazały zniszczenia erozyjnego. Wyniki takich badań, wpływu zawartości niklu stali stopowych na odporność powłok ochronnych na erozję przedstawiono na rys. 5 [12].



Δm – ubytek masy; R_m – wytrzymałość doraźna; δ – wydłużenie względne; Ψ – przewężenie względne.

Rys. 5. Wpływ zawartości niklu na właściwości stali stopowych.

3. Badania porównawcze odporności metali z różną strukturą sieci krystalograficznej i ich wyniki

Na podstawie analizy metod stosowanych w badaniach odporności metali na kawitacyjne zniszczenia erozyjne. [14,15], wybrano metodę kawitacji wentylacyjnej. Ponieważ erozja kawitacyjna jest rezultatem uderzeń mikrostrumieni cieczy po implozji pęcherzyków paro-gazowych, eksperyment przeprowadzono na stanowisku wibracyjnym, z zastosowaniem wody z instalacji miejskiej, jako czynnikiem roboczym.

Weryfikację hipotezy o zmęczeniowym charakterze uszkodzeń erozyjnych wykonano dla próbek z trzech metali, o różnych strukturach sieci krystalicznej: niklu 200/201, stopu niklu monel 400 i stali konstrukcyjnej C45 [5, 6, 7]. Skład chemiczny badanych próbek z niklu 200/201 i monelu 400 podano w tabelach 1 i 2.

Tabela1. Skład chemiczny próbek z niklem 200/201

	C	Si	Mn	S	Co	Cu	Fe	Mg	Ti	Ni	Ni-Co
	0,01	0,04	0,10	<0,01	<0,01	<0,01	<0,01	0,102	0,04	99,67	99,687
Max	0,02	0,15	0,35	0,005	1,0	0,15	0,25	0,15	0,10		
Min											99,6

Tabela 2. Skład chemiczny próbek z monelu 400

		C	Si	Mn	Sr	Al	Co	Cu	Fe	Mg	Ti	Ni	Ni-Co
		0,13	0,23	0,94	0,03	<0,01	0,04	32,6	2,06		0,02	63,9	64,007
Max		0,15	0,5	2,0	0,02	0,5	1,0	34,0	2,5		0,3		
Min								28,0	1,0			36,0	63,0

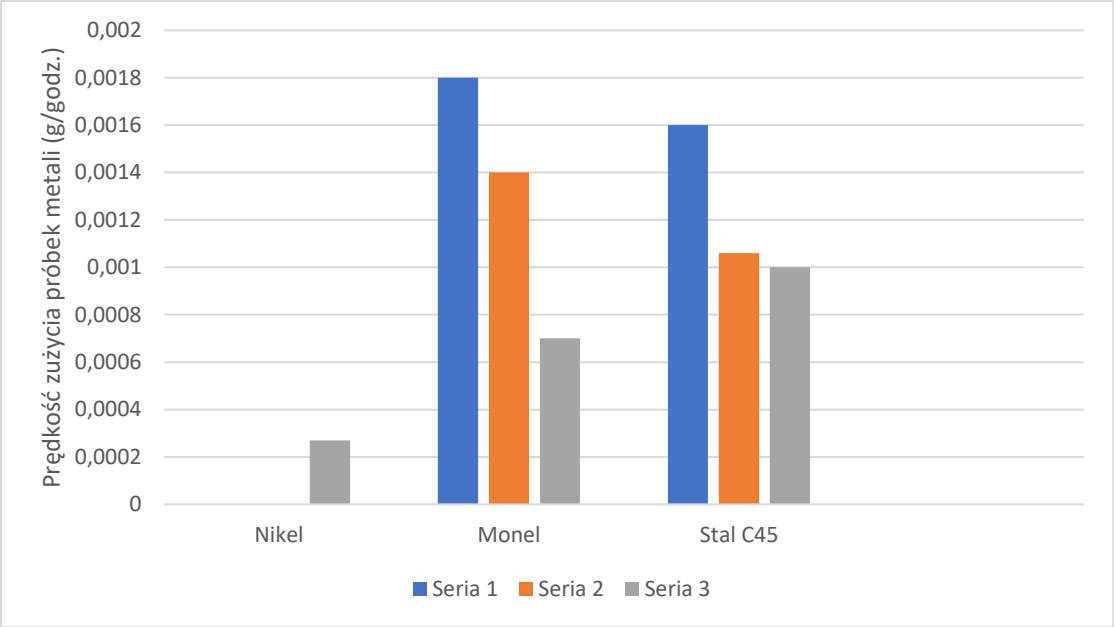
Powierzchnie próbek badanych metali szlifowano do chropowatości odpowiadającej 0,63 mkm, po czym określono twardość ich warstw wierzchnich metodą Vickersa

zgodne z normą PN-EN ISO 6507-1:2000. Wyniki pomiarów jako średnie z sześciu pomiarów, przedstawiono w tabel 3.

Badania przeprowadzono zgodne z normą ASTM G32, w wersji z nieruchomą próbką, w trzech seriach: przez 10 minut odziaływania strumieni na próbki, 30 minut i 60 minut oddziaływania. Po każdej serii próbki ważono na wadze analitycznej. Pomiary ubytku masy próbek dały możliwość określenia prędkości zużycia erozyjnego metali zastosowanych w eksperymencie. Wyniki badań przedstawiono w tabeli 3 i graficznie na rysunku 6.

Tabela 3. Wyniki badań próbek metalowych na stanowisku wibracyjnym

Masa (g)	**Nikel**			**Monel**			**Stal C45**		
	Próbka 1	Próbka 2	Próbka 3	Próbka 1	Próbka 2	Próbka 3	Próbka 1	Próbka 2	Próbka 3
Masa do badań	9,0722	9,1200	8,8327	8,3740	9,7898	9,4627	8,7153	7,9301	7,8240
Masa po 10 min. ekspozycji	9,0722	9,1200	8,8327	8,3739	9,7897	9,4626	8,7150	7,9298	7,8238
Masa po 30 min. ekspozycji	9,072	9,1200	8,8327	8,3738	9,7896	9,4624	8,7147	7,9296	7,8235
Masa po 60 min. ekspozycji	9,0717	9,1199	8,8325	8,3736	9,7894	9,4623	8,7140	7,9292	7,8231
Twardość próbek do badań (metoda) Vickersa	95,8	98,3	112,3	125,0	123,4	125,5	155,6	138,5	126,0
Ubytek masy próbek w trakcie badań	0,0005	0,0001	0,0002	0,0004	0,0003	0,0013	0,0013	0,0009	0,0009
Średni ubytek metalu	0,00027			0,0007			0,001		
Prędkość zużycia po 10 min. ekspozycji (g/godz)	0			0,0018			0,0016		
Prędkość zużycia po 30 min. ekspozycji (g/godz)	0			0,0014			0,00106		
Prędkość zużycia po 60 min. ekspozycji (g/godz)	0,00027			0,0007			0,001		



Rys.6. Prędkość zużycia erozyjnego metali w eksperymencie
Gdzie: seria 1 odpowiada 10 min. ekspozycji badań; seria 2 – 30 min. i seria 3 – 60 min.

4. Podsumowanie

Konfrontując wyniki badań można zauważyć, że najmniejszą odporność na zużycie erozyjne ma stal C45, której prędkość zużycia przekracza prędkość zużycia niklu 200/201 trzydzieści siedem razy. Monel 400 osiąga wartości pośrednie. Fakt ten, jest potwierdzeniem przyjętej hipotezy o skuteczności zastosowania metali plastycznych dla zapobiegania zniszczeniom erozyjny w instalacjach wodnych, podatnych na kawitację. Drugim potwierdzeniem tej hipotezy jest zmniejszenie prędkości zużycia stali C45 i monelu 400 z wzrostem czasu eksperymentu (rys. 6). Jest to skutek zwiększenia twardości warstwy wierzchnich próbek i powstawania zgniotu.

Podsumowując przeprowadzone badania można stwierdzić, że zapobieganie i spowalnianie zużycia erozyjnego powierzchni roboczych elementów maszyn przepływowych może być realizowane poprzez zastosowanie metali plastycznych do ich wytwarzania.

REFERENCES

1. Adamkiewicz A., Waliszyn A.: *Discussion and studies of the properties of a cooling water additive preventing erosive wear of cooled surfaces of ship diesel engines* **.** Eksploatacja i Niezawodność-Maintenance and Reliability, 10/2014, Vol 1,s.565-570, ISSN 1507-2711.
2. Adamkiewicz A., Waliszyn A.: *Studies of erosion resistance of protective coats on the surfaces of machine elements washed with fluids*. Advances in Ma terials Science 6/2018, s. 69-76, ISSN 1730-2439.
3. Amann T., Waidele M., Kailer A., Analysis of mechanical and chemical mechanisms on cavitation erosioncorrosion of steels in salt water using electrochemical

methods, TRIBOLOGY INTERNATIONAL, 2018, Vol. 124: 238-246, DOI: 10.1016/j. triboint.
4. Bolewski Ł., Szkodo M., Kmieć M., Cavitation erosion degradation of Belzona® coatings, Advances in Materials Science, 2017, Vol. 17 (1), DOI: 10.1515/adms-2017-0002.
5. Ciubotariu, CR., Secosan E., Marginean, G., Frunzaverde D., Campian V. C., Experimental Study Regarding the Cavitation and Corrosion Resistance of Stellite 6 and Self-Fluxing Remelted Coatings, STROJNISKI VESTNIK-JOURNAL OF MECHANICAL ENGINEERING, 2016, Vol. 62 (3): 154-162, DOI: 10.5545/sv-jme.2015.2663.
6. Heathcock C.J., Protheroe B.E., Ball A.: Cavitation erosion of stainless steels. Wear, 81(2) (1982) 311-327.
7. Kim J H, Lee M H., A Study on Cavitation Erosion and Corrosion Behavior of Al-, Zn-, Cu-, and Fe-Based Coatings Prepared by Arc Spraying, JOURNAL OF THERMAL SPRAY TECHNOLOGY, 2010, Vol. 19 (6): 1224-1230, DOI: 10.1007/s11666-010-9521-0.
8. Krella A., Cavitation degradation model of hard thin PVD coatings, Advances in Materials Science, 2010, Vol. 10 (3): 27–36, DOI: 10.2478/v10077-010-0010-4
9. Kumar, H., Chittosiya, C., Shukla, V.N.: HVOF Sprayed WC Based Cermet Coating for Mitigation of Cavitation, Erosion & Abrasion in Hydro Turbine Blade, MATERIALS TODAY-PROCEEDINGS, 2018, Vol. 5 (2) : 6413-6420.
10. Krumenacker, L., Fortes-Patella, R., Archer, A., Numerical estimation of cavitation intensity, IOP Conference Series-Earth and Environmental Science, 2014, Vol. 22, Article Number: UNSP 052014, DOI: 10.1088/1755-1315/22/5/052014.
11. Kwok C. T., Man H. C., Cheng F. T.: Cavitation erosion and damage mechanisms of alloys with duplex structures. Materials Science and Engineering A242 (1998) 108-120.
12. Steller J., Krella A., Koronowicz J., Janicki W.: Towards quantitative assessment of material resistance to cavitation erosion. Wear, 258 (2005) 604–613.
13. Waliszyn A., Adamkiewicz A.: A metod of vibration damping for diesel engeene cylinder lines to prevent the consequences of erosion. Eksploatacja I Niezawodność - Maintenance and Reliability, 2018, Vol. 20, s. 371-377, ISSN 1507-2711.
14. Yang D., Yu A., Ji B., Zhou, J., Luo X., Numerical analyses of ventilated cavitation over a 2-D NACA0015 hydrofoil using two turbulence modeling methods, JOURNAL OF HYDRODYNAMICS, 2018, Vol. 30 (2): 345-356, DOI: 10.1007/s42241-018-0032-7.
15. Yu A., Luo X., Ji B., Analysis of ventilated cavitation around a cylinder vehicle with nature cavitation using a new simulation method, SCIENCE BULLETIN, 2015, Vol. 60 (21): 1833-1839, DOI: 10.1007/s11434-015-0916-7.