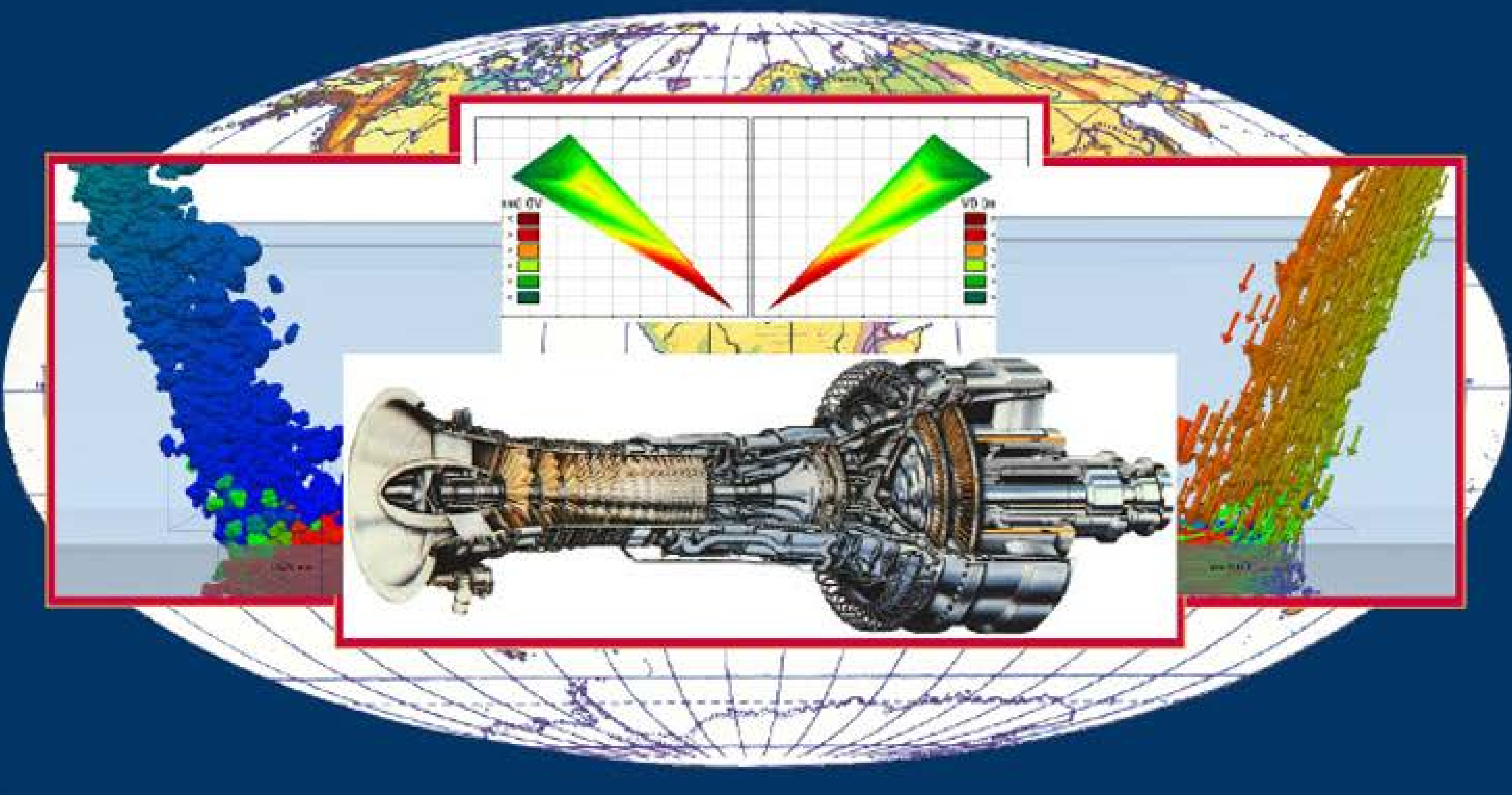
Vol. 22. No 3, 2020

ISSN 1507-2711
Cena: 25 zł

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY



Polskie Naukowo Techniczne Towarzystwo Eksploatacyjne
Warszawa

Polish Maintenance Society
Warsaw

LIU D, WANG S, TOMOVIC MM. **Degradation modeling method for rotary lip seal based on failure mechanism analysis and stochastic process**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 381–390, http://dx.doi.org/10.17531/ein.2020.3.1.

Rotary lip seal is widely used in aircraft and its performance affects the safety of the aircraft. Hence, it is necessary to estimate useful lifetime and reliability of the seal. Degradation of rotary lip seal is always with random effects, which cannot be considered by theoretical failure mechanism analysis. Hence, in order to consider the random effects of rotary lip seal degradation, stochastic processes are applied. Furthermore, considering the monotonic degradation of the seal, Gamma process and inverse Gaussian process are selected as the candidate processes. To combine the candidate processes, Bayesian model averaging is introduced. Based on the failure mechanism analysis and numerical simulation, the theoretical wear path is predicted and corresponding linearization method is proposed. The measured degradation data is converted and the seal wear process is transformed to a linear degradation process. The model parameters and model probabilities are evaluated by fully Bayesian inference method. The effectiveness of the proposed method is verified by comparing the predicting degradation and experimental observations. The proposed method can be used to evaluate reliability and useful lifetime of rotary lip seal. According to sensitivity analysis, an effective way to improve lifetime and reliability of the seal is to increase the wear depth threshold.

PAWEŁCZYK M, FULARA S, SEPE M, DE LUCA A, BADORA M. **Industrial gas turbine operating parameters monitoring and data-driven prediction**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 391–399, http://dx.doi.org/10.17531/ein.2020.3.2.

The article reviews traditional and modern methods for prediction of gas turbine operating characteristics and its potential failures. Moreover, a comparison of Machine Learning based prediction models, including Artificial Neural Networks (ANN), is presented. The research focuses on High Pressure Compressor (HPC) recoup pressure level of 4th generation LM2500 gas generator (LM2500+G4) coupled with a 2-stage High Speed Power Turbine Module. The researched parameter is adjustable and may be used to balance net axial loads exerted on thrust bearing to ensure stable gas turbine operation, but its direct measurement is technically difficult implicating the need to indirect measurement via set of other gas turbine sensors. Input data for the research have been obtained from BHGE manufactured and monitored gas turbines and consists of real-time data extracted from industrial installations. Machine learning models trained using the data show less than 1% Mean Absolute Percentage Error (MAPE) as obtained with the use of Random Forest and Gradient Boosting Regression models. Multilayer Perceptron Artificial Neural Networks (MLP ANN) models are reviewed, and their performance checks inferior to Random Forest algorithm-based model. The importance of hyperparameter tuning and feature engineering is discussed.

ÖZCAN E, DANIŞAN T, YUMUŞAK R, EREN T. **An artificial neural network model supported with multi criteria decision making approaches for maintenance planning in hydroelectric power plants**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 400–418, http://dx.doi.org/10.17531/ein.2020.3.3.

Power plants are the large-scale production facilities with the main purpose of realizing uninterrupted, reliable, efficient, economic and environmentally friendly energy generation. Maintenance is one of the critical factors in achieving these comprehensive goals, which are called as sustainable energy supply. The maintenance processes carried out in order to ensure sustainable energy supply in the power plants should be managed due to the costs arising from time requirement, the use of material and labor, and the loss of generation. In this respect, it is critical that the fault dates are forecasted, and maintenance is performed without failure in power plants consisting of thousands of equipment. In this context in this study, the maintenance planning problem for equipment with high criticality level is handled in one of the large-scale hydroelectric power plants that meet the quintile of Turkey's energy demand as of the end of 2018. In the first stage, the evaluation criteria determined by the power plant experts are weighted by the Analytical Hierarchy Process (AHP), which is an accepted method in the literature, in order to determine the criticality levels of the equipment in terms of power plant at the next stage. In order to obtain the final priority ranking of the equipment in terms of power plant within the scope of these weights, Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is used because of its advantages compared to other outranking algorithms. As a result of this solution, for the 14 main equipment groups with the highest criticality level determined on the basis of the power plant, periods between two breakdowns are estimated, and maintenance planning is performed based on these periods. In the estimation phase, an artificial neural network (ANN) model has been established by using 11-years fault data for selected equipment groups and the probable fault dates are estimated by considering a production facility as a system without considering

LIU D, WANG S, TOMOVIC MM. **Metoda modelowania degradacji obrotowego uszczelnienia wargowego w oparciu o analizę mechanizmu uszkodzenia i proces stochastyczny**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 381–390, http://dx.doi.org/10.17531/ein.2020.3.1.

Obrotowe uszczelnienia wargowe znajdują szerokie zastosowanie w samolotach, a ich sprawność wpływa na bezpieczeństwo statków powietrznych. Oznacza to, iż szacowanie żywotności i niezawodności tego rodzaju uszczelnień ma kluczowe znaczenie. Degradacja obrotowego uszczelnienia wargowego jest zawsze związana z efektami losowymi, których nie uwzględnia teoretyczna analiza mechanizmu uszkodzenia. Dlatego też do oceny efektów losowych degradacji obrotowego uszczelnienia wargowego wykorzystuje się procesy stochastyczne, takie jak proces Gamma czy odwrotny proces Gaussa. W przedstawionej pracy, wybrane procesy degradacji łączono za pomocą metody bayesowskiego uśredniania modeli. Na podstawie analizy mechanizmów uszkodzeń i symulacji numerycznej, konwertowano uzyskane w pomiarach dane degradacyjne, co pozwoliło na przekształcenie procesu degradacji obrotowego uszczelnienia wargowego w proces liniowy. Parametry modelu i prawdopodobieństwa oceniano za pomocą metody pełnego wnioskowania bayesowskiego na podstawie obserwacji degradacji. Skuteczność przedstawionej metody weryfikowano porównując przewidywane i obserwowane wartości degradacji. Proponowaną metodę można wykorzystywać do oceny niezawodności i żywotności obrotowego uszczelnienia wargowego. Przeprowadzona analiza czułości pokazuje, że skutecznym sposobem na poprawę żywotności i niezawodności omawianego typu uszczelnienia jest zwiększenie progu uszkodzenia w postaci maksymalnej głębokości zużycia.

PAWEŁCZYK M, FULARA S, SEPE M, DE LUCA A, BADORA M. **Monitorowanie oraz bazująca na danych predykcja parametrów roboczych przemysłowej turbiny gazowej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 391–399, http://dx.doi.org/10.17531/ein.2020.3.2.

W artykule przedstawiono przegląd klasycznych i aktualnych metod przewidywania parametrów operacyjnych oraz potencjalnych usterek turbin gazowych. Dodatkowo zaprezentowano porównanie wybranych modeli opartych o uczenie maszynowe, w tym modeli wykorzystujące sztuczne sieci neuronowe. Przeprowadzone badania dotyczyły analiz poziomu ciśnienia ze sprężarki turbiny gazowej LM2500 czwartej generacji (LM2500+G4) połączonej z dwustopniową turbiną roboczą. Badany parametr podlega sterowaniu i może posłużyć do wyrównania sił osiowych działających na łożysko główne wału wysokiego ciśnienia w celu zapewnienia stabilnej i niezawodnej pracy turbiny gazowej. Jednocześnie jego bezpośredni pomiar jest kosztowny stąd potrzeba dokonania pośredniego pomiaru z wykorzystaniem innych czujników zamontowanych na turbinie. Dane wejściowe do analiz otrzymano dzięki uprzejmości producenta turbin, firmy BHGE. Zawierają one parametry bezpośrednio pobrane z monitorowanych turbin gazowych. Modele uczenia maszynowego otrzymane w wyniku analizy charakteryzują się średnim błędem procentowym (MAPE) na poziomie poniżej 1%. Najmniejszym błędem charakteryzują się modele otrzymane przy zastosowaniu lasów losowych (Random Forest) oraz gradientowego wzmacniania regresji (Gradient Boosted Regression). Przetestowano także zastosowanie wielowarstwowych, w pełni połączonych sztucznych sieci neuronowych, których efektywność okazała się niższa od modelu opartego o algorytm lasów losowych. W podsumowaniu podkreślono wagę dostosowywania hiperparametrów i inżynierii cech.

ÖZCAN E, DANIŞAN T, YUMUŞAK R, EREN T. **Planowanie utrzymania ruchu w elektrowniach wodnych w oparciu o model sztucznej sieci neuronowej wsparty wielokryterialnymi metodami podejmowania decyzji**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 400–418, http://dx.doi.org/10.17531/ein.2020.3.3.

Elektrownie to zakłady produkcyjne o dużej skali, których głównym celem jest nieprzerwane, niezawodne, wydajne, rentowne oraz przyjazne dla środowiska wytwarzanie energii. Utrzymanie ruchu stanowi jeden z kluczowych czynników pozwalających na osiągnięcie tych szeroko zakrojonych celów, które określa się wspólnym mianem zrównoważonych dostaw energii. W elektrowniach, procesami utrzymania ruchu, realizowanymi w celu zapewnienia zrównoważonych dostaw energii, zarządza się z uwzględnieniem kosztów związanych z wymogami czasowymi, kosztów materiałów i robocizny oraz strat wytwarzania energii. Ponieważ elektrownie wykorzystują tysiące różnych urządzeń, niezwykle ważne jest prognozowanie dat wystąpienia uszkodzeń oraz zapewnienie bezawaryjnego utrzymania ruchu. W przedstawionych badaniach, rozważano problem planowania utrzymania ruchu sprzętu o wysokim poziomie krytyczności na przykładzie jednej z dużych elektrowni wodnych, która na koniec 2018 r. pokrywała jedną piątą zapotrzebowania Turcji na energię elektryczną. W pierwszym etapie badań, kryteria oceny określone przez ekspertów zatrudnionych w elektrowni ważono za pomocą powszechnie stosowanej w literaturze metody procesu hierarchii analitycznej (AHP) w celu ustalenia poziomów krytyczności poszczególnych elementów wyposażenia elektrowni. Aby opracować ostateczny ranking priorytetowości elementów wyposażenia elektrowni na podstawie określonych wcześniej wag, zastosowano technikę TOPSIS, która polega na porządkowaniu preferencji na podstawie podobieństwa do idealnego rozwiązania. Techniki tej użyto ze względu na jej zalety, których nie mają inne algorytmy oparte na relacji przewyższania (ang. outranking algorithms). Na podstawie wyników otrzymanych dla 14 głównych grup urządzeń o najwyższym poziomie krytyczności, określonym na podstawie danych pochodzących z elektrowni, oszacowano czasy pomiędzy dwiema awariami, a na ich podstawie zaplanowano działania konserwacyjne. W fazie szacowania, opracowano model the sector for the first time in the literature. With the plan including the maintenance activities that will be carried out before the determined breakdown dates, increasing the generation efficiency, extending the economic life of the power plant, minimizing the generation costs, maximizing the plant availability rate and maximizing profit are aimed. The maintenance plan is implemented for 2 years in the power plant and the unit shutdowns resulting from the selected equipment groups are not met and the mentioned goals are reached.

sztucznej sieci neuronowej (ANN) w oparciu o dane o uszkodzeniach, które wystąpiły w ostatnich 11 latach działania elektrowni, dla wybranych grup urządzeń. Przewidywane daty wystąpienia uszkodzeń szacowano, po raz pierwszy w literaturze, biorąc pod uwagę zakład produkcyjny jako system, bez uwzględnienia sektora produkcyjnego. Plan obejmuje działania konserwacyjne, które mają być przeprowadzone przed przewidywanymi datami awarii, w celu zwiększenia wydajności wytwarzania energii, przedłużenia żywotności elektrowni, minimalizacji kosztów wytwarzania energii, maksymalizacji wskaźnika dostępności elektrowni oraz maksymalizacji zysków. Opracowany plan konserwacji wdrażano w omawianej elektrowni przez 2 lata. W tym okresie nie odnotowano przerw w pracy jednostek wytwórczych spowodowanych awarią rozważanych grup urządzeń, co oznacza, że wspomniane cele zostały osiągnięte.

SHAO Y, LIU B, WANG S, XIAO P. **A novel test case prioritization method based on problems of numerical software code statement defect prediction**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 419–431, http://dx.doi.org/10.17531/ein.2020.3.4.

Test case prioritization (TCP) has been considerably utilized to arrange the implementation order of test cases, which contributes to improve the efficiency and resource allocation of software regression testing. Traditional coverage-based TCP techniques, such as statement-level, method/function-level and class-level, only leverages program code coverage to prioritize test cases without considering the probable distribution of defects. However, software defect data tends to be imbalanced following Pareto principle. Instinctively, the more vulnerable the code covered by the test case is, the higher the priority it is. Besides, statement-level coverage is a more fine-grained method than function-level coverage or class-level coverage, which can more accurately formulate test strategies. Therefore, we present a test case prioritization approach based on statement software defect prediction to tame the limitations of current coverage-based techniques in this paper. Statement metrics in the source code are extracted and data pre-processing is implemented to train the defect predictor. And then the defect detection rate of test cases is calculated by combining the prioritization strategy and prediction results. Finally, the prioritization performance is evaluated in terms of average percentage faults detected in four open source datasets. We comprehensively compare the performance of the proposed method under different prioritization strategies and predictors. The experimental results show it is a promising technique to improve the prevailing coverage-based TCP methods by incorporating statement-level defect-proneness. Moreover, it is also concluded that the performance of the additional strategy is better than that of max and total, and the choice of the defect predictor affects the efficiency of the strategy.

SHAO Y, LIU B, WANG S, XIAO P. **Nowatorska metoda priorytetyzacji przypadków testowych oparta na prognozowaniu błędów instrukcji kodu oprogramowania numerycznego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 419–431, http://dx.doi.org/10.17531/ein.2020.3.4.

Metodę priorytetyzacji przypadków testowych (TCP) wykorzystuje się powszechnie do ustalania kolejności implementacji przypadków testowych, co przyczynia się do poprawy wydajności i alokacji zasobów w trakcie testowania regresyjnego oprogramowania. Tradycyjne techniki TCP oparte na pokryciu na poziomie instrukcji, metody/funkcji oraz klasy, wykorzystują pokrycie kodu programu tylko w celu ustalenia priorytetów przypadków testowych, bez uwzględnienia prawdopodobnego rozkładu błędów. Jednak dane o błędach oprogramowania są zwykle niezrównoważone zgodnie z zasadą Pareto. Instynktownie, im bardziej wrażliwy jest kod pokryty przypadkiem testowym, tym wyższy jest jego priorytet. Poza tym, pokrycie na poziomie instrukcji jest bardziej szczegółową metodą niż pokrycie na poziomie funkcji lub pokrycie na poziomie klasy, które mogą dokładniej formułować strategie testowe. Dlatego w artykule przedstawiamy podejście do priorytetyzacji przypadków testowych oparte na prognozowaniu błędów instrukcji oprogramowania, które pozwala zmniejszyć ograniczenia obecnych technik opartych na pokryciu. Wyodrębniono metryki instrukcji w kodzie źródłowym i zaimplementowano wstępne przetwarzanie danych w celu nauczania predyktora błędów. Następnie obliczono wskaźnik wykrywania błędów w przypadkach testowych poprzez połączenie strategii priorytetyzacji i wyników prognozowania. Wreszcie, oceniono wydajność ustalania priorytetów pod względem średnich procentowych błędów wykrytych w czterech zestawach danych typu open source. Kompleksowo porównano wydajność proponowanej metody w ramach różnych strategii ustalania priorytetów i predyktorów. Wyniki eksperymentów pokazują, że jest to obiecująca technika poprawy dominujących metod TCP opartych na pokryciu poprzez włączenie podatności na błędy na poziomie instrukcji. Ponadto stwierdzono również, że strategia dodatkowa cechuje się lepszą wydajnością niż strategie max i total, a wybór predyktora błędów wpływa na skuteczność strategii.

DZIERWA A, GAŁDA L, TUPAJ M, DUDEK K. **Investigation of wear resistance of selected materials after slide burnishing process**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 432–439, http://dx.doi.org/10.17531/ein.2020.3.5.

The article presents the research on the impact of slide burnishing process carried out with use of various ceramics on friction and wear of steel elements. In addition, surfaces after grinding, lapping and polishing processes were tested. The tribological couple was made of steel discs, toughened to a hardness of 40 ± 2 HRC, and balls made of 100Cr6 steel with a hardness of 62 HRC. The tests were carried out at three sliding speeds: 0.16 m/s, 0.32 m/s and 0.48 m/s. The research proved the possibility of improving selected tribological properties of friction pairs thanks to the use of slide burnishing process and also allowed to establish a number of relationships between the parameters characterizing the surface topography and the tribological parameters.

DZIERWA A, GAŁDA L, TUPAJ M, DUDEK K. **Badania odporności na zużycie wybranych materiałów poddanych procesowi nagniatania ślizgowego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 432–439, http://dx.doi.org/10.17531/ein.2020.3.5.

W artykule przedstawiono wyniki badań wpływu procesu nagniatania ślizgowego realizowanego z wykorzystaniem różnych ceramik na wielkość zużycia oraz siłę tarcia elementów stalowych. Dodatkowo badaniom poddano powierzchnie po procesach szlifowania, docierania oraz polerowania. Skojarzenie materiałowe stanowiły tarcze stalowe ulepszone cieplnie do twardości 40±2 HRC oraz kulki ze stali100Cr6 o twardości 62 HRC. Badania zrealizowano przy trzech prędkościach poślizgu: 0,16 m/s, 0,32 m/s oraz 0,48 m/s. Badania udowodniły możliwość poprawy wybranych właściwości tribologicznych par trących dzięki zastosowaniu procesu nagniatania ślizgowego a także pozwoliły na ustalenie szeregu zależności pomiędzy parametrami charakteryzującymi strukturę geometryczną powierzchni oraz parametrami tribologicznymi.

RODRIGUES J, COSTA I, TORRES FARINHA J, MENDES M, MARGALHO L. **Predicting motor oil condition using artificial neural networks and principal component analysis**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 440–448, http://dx.doi.org/10.17531/ein.2020.3.6.

The safety and performance of engines such as Diesel, gas or even wind turbines depends on the quality and condition of the lubricant oil. Assessment of engine oil condition is done based on more than twenty variables that have, individually, variations that depend on the engines' behaviour, type and other factors. The present paper describes a model to automatically classify the oil condition, using Artificial Neural Networks and Principal Component Analysis. The study was done using data obtained from two passenger bus companies in a country of Southern Europe. The results show the importance of each variable monitored for determining the ideal time to change oil. In many cases, it may be possible to enlarge intervals between maintenance interventions, while in other cases the oil passed the ideal change point.

RODRIGUES J, COSTA I, TORRES FARINHA J, MENDES M, MARGALHO L. **Prognozowanie stanu oleju silnikowego za pomocą sztucznych sieci neuronowych i analizy składowych głównych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 440–448, http://dx.doi.org/10.17531/ein.2020.3.6.

Bezpieczeństwo i wydajność silników takich, jak silniki Diesla czy gazowe, a nawet turbiny wiatrowe, zależą od jakości i stanu oleju smarowego. Stanu oleju silnikowego ocenia się na podstawie ponad dwudziestu zmiennych, z których każda ulega wahaniom w zależności od typu i zachowania silnika oraz innych czynników. W niniejszym artykule opisano model, który pozwala na automatyczną klasyfikację stanu oleju, z wykorzystaniem sztucznych sieci neuronowych i analizy składowych głównych. Badania przeprowadzono na podstawie danych uzyskanych od dwóch przewoźników pasażerskich działających na terenie jednego z krajów położonych na południu Europy. Wyniki pokazują, że każda z monitorowanych zmiennych ma znaczenie dla określenia idealnego czasu na wymianę oleju. Podczas gdy w wielu przypadkach w badanych przedsiębiorstwach możliwe było zwiększenie odstępów czasowych między działaniami konserwacyjnymi, w innych, idealny moment wymiany oleju został przekroczony.

WIĘCŁAWSKI K, MĄCZAK J, SZCZUROWSKI K. **Electric current as a source of information about control parameters of indirect injection fuel injector**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 449–454, http://dx.doi.org/10.17531/ein.2020.3.7.

The article discusses results of the laboratory experiments in which fuel injectors used in indirect injection internal combustion engines were tested. During the experiments, numerous dosing cycles of the injectors were performed while changing the control parameters, due to which, the dosing characteristics were developed and influence of applied parameters on the resultant fuel flow determined. Simultaneously, the voltage and electric current waveforms in the injector coil were recorded, due to which finding links between the electric current characteristics and the determinants of the injector work was possible. The investigation has shown that parameters of electric current constitute a precise criterion for assessing the operation of the solenoid valve, because fuel flow is created due to the work ofelectric current. Thus, by observing the changes in the current flowing through the valve coil,it is possible to monitor precisely the correctness of the process of opening the flow and the electric current intensity, at which the flow began and to determine the mechanical quantities such as fuel dose and pressure. As a result, a characteristic is developed, that provides the links between the fuel pressure and the electric current at the point of lifting the needle, which is quite a novel approach. Such a characteristic can be used in diagnostics and control of fuel injectors as well as all kinds of electromagnetic valves.

WIĘCŁAWSKI K, MĄCZAK J, SZCZUROWSKI K. **Przebieg prądowy jako źródło informacji o parametrach sterowania wtryskiwaczem paliwowym wtrysku pośredniego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 449–454, http://dx.doi.org/10.17531/ein.2020.3.7.

Artykuł przedstawia wyniki eksperymentów laboratoryjnych polegających na testowaniu wtryskiwaczy paliwowych stosowanych w silnikach spalinowych z wtryskiem pośrednim. Podczas eksperymentów wykonano wiele cykli dawkowania wtryskiwaczy zmieniając parametry sterowania, dzięki czemu opracowano charakterystyki dawkowania i określono wpływ stosowanych parametrów sterowania na wynikowy przepływ paliwa. Jednocześnie rejestrowano przebiegi napięcia i natężenia prądu elektrycznego w cewce wtryskiwacza, dzięki czemu możliwe było powiązanie charakterystyk prądowych z determinantami pracy wtryskiwacza. Wykazano, iż parametry prądowe są precyzyjnym kryterium oceny pracy zaworu elektromagnetycznego, ponieważ dzięki wykonanej przez prąd pracy powstaje przepływ paliwa. Zatem poprzez obserwację zmian prądu płynącego przez cewkę zaworu, można precyzyjnie monitorować prawidłowość procesu otwierania przepływu oraz natężenie prądu, przy którym przepływ się rozpoczął oraz określać wielkości mechaniczne jak dawka i ciśnienie paliwa. Wynikiem badań jest opracowanie charakterystyki wiążącej ciśnienie paliwa z natężeniem prądu w punkcie podnoszenia iglicy, co jest podejściem nowatorskim. Taka charakterystyka może być wykorzystana w diagnostyce i sterowaniu wtryskiwaczy paliwowych oraz wszelkiego rodzaju zaworów elektromagnetycznych.

YAN S, MA B, WANG X, CHEN J, ZHENG C. **Maintenance policy for oil-lubricated systems with oil analysis data**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 455–464, http://dx.doi.org/10.17531/ein.2020.3.8.

Maintenance of oil-lubricated systems plays a significant role in reducing unexpected system failures and improving machine availability. This paper deals with the oil-lubricated systems subject to gradual degradation that revealed by metal wear debris monitored using oil analysis. Oil-lubricated systems usually undertake several preventive maintenances during operation, after each maintenance, the system typical restores to an intermediate state between good-as-new state and bad-as-old state due to system aging such as cumulative wear. Furthermore, oil-lubricated systems often operate continuously in mission execution with availability constraints. However, existing literature still lacks a method to integrate the availability constraints with the system aging into the cause of optimizing the maintenance policy. To fill this gap, this paper develops a maintenance policy optimization method to determine the optimal maintenance threshold joint considering the availability constraints and the system aging. A case study of the power-shift steering transmission systems modelled by a wiener process is presented to illustrate the proposed method in practical application.

YAN S, MA B, WANG X, CHEN J, ZHENG C. **Polityka utrzymania ruchu układów smarowanych olejem w oparciu o dane z analizy oleju**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 455–464, http://dx.doi.org/10.17531/ein.2020.3.8.

Konserwacja układów smarowanych olejem odgrywa istotną rolę w eksploatacji maszyn, umożliwiając zmniejszenie liczby nieoczekiwanych uszkodzeń i poprawiając dostępność maszyn. Niniejszy artykuł dotyczy układów smarowanych olejem ulegających stopniowej degradacji, którą można mierzyć za pomocą analizy oleju, monitorując zawartość drobin metalu powstających na skutek zużycia. Podczas swojej pracy, układy smarowane olejem zwykle poddawane są kilkakrotnie przeglądom zapobiegawczym. Po każdej konserwacji, układ wraca do stanu pośredniego między stanem fabrycznej nowości (as good as new) a stanem "jak przed konserwacją" (as bad as old), co wynika ze starzenia się systemu, m.in. skumulowanego zużycia. Co więcej, systemy smarowane olejem często działają w sposób ciągły, wykonując misje z ograniczeniami dostępności. Jednak w istniejącej literaturze wciąż brakuje metody, która pozwalałaby na zintegrowane ujęcie ograniczeń dostępności i starzenia w celu optymalizacji polityki utrzymania ruchu. Aby wypełnić tę lukę, w niniejszym artykule opracowano metodę optymalizacji polityki utrzymania ruchu, dzięki której można określić optymalny próg konserwacji z uwzględnieniem zarówno ograniczeń dostępności jak i starzenia się systemu. Możliwość praktycznego zastosowania zaproponowanej metody zilustrowano na podstawie studium przypadku układów przekładni kierowniczych zamodelowanych za pomocą procesu Wienera.

KOWALCZYK J, MADEJ M, OZIMINA D. **Evaluation of performance characteristics of the environmentally friendly cutting fluid with zinc aspartate**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 465–471, http://dx.doi.org/10.17531/ein.2020.3.9.

The effect of the cutting fluid with zinc aspartate on the quality of the workpiece surface layer is reported. Until now, zinc aspartate has been used primarily in medicine and pharmacology. This paper compares the ecological cutting fluid containing zinc aspartate with a classic mineral oil-based coolant. Toxicity tests and a controlled process of tool wear during face turning were performed. Test results indicate that the use of zinc aspartate-based cutting fluids contributes to the reduction of the material roughness parameter values up to 35%, benefitting the final quality of the workpiece.

KOWALCZYK J, MADEJ M, OZIMINA D. **Ocena właściwości eksploatacyjnych proekologicznej cieczy chłodząco-smarującej zawierającej asparaginian cynku**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 465–471, http://dx.doi.org/10.17531/ein.2020.3.9.

W pracy przedstawiono wyniki badań wpływu cieczy chłodząco-smarującej z asparaginianem cynku na jakość technologiczną warstwy wierzchniej obrabianych elementów. Asparaginian cynku dotychczas nie był stosowany w takich rozwiązaniach, głównie wykorzystywany był w medycynie i farmakologii. W badaniach przeprowadzono analizę porównawczą proekologicznego chłodziwa zawierającego asparaginian cynku z klasycznym chłodziwem opartym na bazie oleju mineralnego. Ciecze chłodząco-smarujące poddano badaniom toksyczności oraz wykonano kontrolowany proces eksploatacji narzędzi w czasie toczenia poprzecznego. Wyniki badań wskazują, że zastosowanie chłodzenia cieczą na bazie asparaginianu cynku redukuje parametry chropowatości obrabianego materiału nawet o 35%, korzystnie wpływając na jakość finalną detalu.

ANDREJIOVA M, GRINCOVA A, MARASOVA D. **Analysis of tensile properties of worn fabric conveyor belts with renovated cover and with the different carcass type**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 472–481, http://dx.doi.org/10.17531/ein.2020.3.10.

Conveyors are the means of transportation used in many industries. The load-bearing and tractive component of a belt conveyor is a conveyor belt which consists of a carcass and cover layers. During an operation, belts are exposed to loads that cause damage to the belts. It is therefore necessary to ensure that a conveyor belt possesses required mechanical properties during the transport of material. The key mechanical properties of a conveyor belt are tensile properties. They are significantly affected by the fabric carcass of a conveyor belt. The tensile properties of conveyor belts are largely affected by the carcass materials. They are also affected by the types of fibres in the longitudinal (warp) and transverse (weft) directions of the fabric carcass because the carcass transfers all tensile stresses of the conveyor belt. A fabric conveyor belt is regarded as a composite material, consisting of the carcass (polyamide P and polyamide-polyester EP) and the cover layers. The costs of a conveyor belt represent 10-30 % of the price of the entire conveyor. It is therefore reasonable to prefer only those

ANDREJIOVA M, GRINCOVA A, MARASOVA D. **Analiza właściwości wytrzymałościowych zużytych tkaninowych taśm przenośnikowych z różnymi typami rdzenia po renowacji górnej okładki**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 472–481, http://dx.doi.org/10.17531/ein.2020.3.10.

Przenośniki taśmowe znajdują zastosowanie jako urządzenia transportowe w wielu gałęziach przemysłu. Nośnym i pociągowym elementem przenośnika taśmowego jest taśma transportowa, która zbudowana jest z rdzenia i okładek. Podczas pracy, taśmy narażone są na obciążenia, które prowadzą do ich uszkodzenia. Dlatego konieczne jest aby taśma transportująca materiał posiadała wymagane właściwości mechaniczne. Kluczowymi właściwościami mechanicznymi taśmy przenośnikowej są właściwości wytrzymałościowe. Zależą one w dużym stopniu od tkaninowego rdzenia taśmy, a w szczególności od rodzaju materiałów, z których jest zbudowany, oraz typu włókien wchodzących w skład jego osnowy (biegnących w kierunku wzdłużnym) i wątku (w kierunku poprzecznym), jako że to właśnie rdzeń przenosi wszystkie naprężenia rozciągające taśmy przenośnika. Tkaninową taśmę transportową uważa się za materiał kompozytowy, składający się z rdzenia (poliamid P i poliamid–poliester EP) oraz warstw wierzchnich (okładek). Koszty taśmy stanowią conveyor belts that show the properties prescribed by relevant norms. The subject of the article is worn conveyor belts with renovated top cover (renovated conveyor belts). The tensile properties are used to assess the suitability for further use of renovated conveyor belts in practice. The article presents the analysis of tensile properties of renovated fabric conveyor belts in relation to the carcass type. The observed results were compared applying the DOE method, regression and correlation analysis, and the method of statistic induction. All the conclusions, made based on the above-listed methods, are identical for the examined tensile properties. The results indicate that the examined tensile properties of conveyor belts have not undergone any significant change after the renovation of the top cover layer.

10–30% ceny całego przenośnika, dlatego do użycia powinno dopuszczać się jedynie takie taśmy przenośnikowe, które wykazują właściwości określone w odpowiednich normach. Przedmiotem artykułu są zużyte taśmy przenośnikowe z odnowioną górną okładką (odnowione taśmy przenośnikowe). Właściwości wytrzymałościowe wykorzystano do oceny przydatności odnowionych taśm do dalszego wykorzystania w warunkach praktycznych. W artykule przedstawiono analizę właściwości wytrzymałościowych odnowionych taśm tkaninowych w funkcji typu rdzenia. Zaobserwowane wyniki porównano stosując metodę DOE, analizę regresji i korelacji oraz metodę indukcji statystycznej. Wszystkie wnioski uzyskane w oparciu o wyżej wymienione metody są identyczne dla badanych właściwości wytrzymałościowych. Wyniki wskazują, że badane właściwości wytrzymałościowe taśm przenośnikowych nie ulegają istotnej zmianie po renowacji górnej okładki.

STETTER R, GÖSER R, GRESSER S, TILL M, WITCZAK M. **Fault-tolerant design for increasing the reliability of an autonomous driving gear shifting system**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 482–492, http://dx.doi.org/10.17531/ein.2020.3.11.

The reliability of technical systems can be greatly reduced if possible faults cannot be accommodated but lead to system shut-down with sometimes catastrophic consequences. The algorithms and systems of fault-tolerant control were developed in the last years into a powerful tool to accommodate such faults. Additionally, it became obvious that the design of a technical system can ease or hinder the application of these tools and can also lead to the accommodation of faults be itself. This kind of design – fault-tolerant design – and its components are presented in this paper on the example of a shifting system for the gear box an autonomous driving race car. This race car competes in the well-known formula student driverless competition; in such competitions the reliability of the car and the capability to accommodate not avoidable faults is of paramount importance. The different elements of fault-tolerance incorporated in the design of the gear shifting system are explained on the basis of an established model of product concretization.

STETTER R, GÖSER R, GRESSER S, TILL M, WITCZAK M. **Projektowanie tolerujące uszkodzenia zwiększające niezawodności systemu zmiany biegów pojazdu autonomicznego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 482–492, http://dx.doi.org/10.17531/ein.2020.3.11.

Niezawodność systemów technicznych może być znacznie ograniczona w przypadku braku odpowiedniej akomodacji uszkodzeń, która może doprowadzić do awarii systemu mogącej mieć katastrofalne konsekwencje. W celu przeciwdziałania temu niepożądanemu zjawisku, w ostatnich latach opracowano szereg algorytmów sterowania tolerującego uszkodzenia, umożliwiających odpowiednią akomodację uszkodzeń. Dodatkowo, oczywistym jest, że sposób projektowania danego systemu może ułatwić lub utrudnić funkcjonowanie powyższych algorytmów. Może one również sam w sobie umożliwiać odpowiednią akomodację uszkodzeń. Tak sposób projektowania, projektowanie tolerujące uszkodzenia, jest przedmiotem niniejszej pracy na przykładzie systemu zmiany biegów w autonomicznych pojeździe wyścigowym. Powyższy pojazd współzawodniczy w znanych studenckich zawodach wyścigowych pojazdów autonomicznych. Oczywistym jest fakt, że w tego typu zawodach, niezawodność pojazdu i jego zdolność akomodacji uszkodzeń jest szczególnie ważna. W pracy rozważa się różne element projektowania tolerującego uszkodzenia systemu zmiany biegów opisanego na podstawie ustalonego modelu konkretyzacji produktu.

ŠKERLIČ S, SOKOLOVSKIJ E, ERČULJ V. **Maintenance of heavy trucks: an international study on truck drivers**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 493–500, http://dx.doi.org/10.17531/ein.2020.3.12.

Since the implementation of modern approaches in operation maintenance, drivers are expected to be integrated into the entire system of maintenance in order to take over the professional competencies of maintenance workers. For this purpose, an international study was conducted on a sample of 249 truck drivers, with the aim to determine how maintenance in transport companies affects the role of heavy truck drivers in fleet maintenance. Based on the developed SEM model, it was determined that building an efficient maintenance infrastructure in transport companies enables the active participation of truck drivers in the company's maintenance system. Drivers are capable of repairing minor failures, but only if they are proactive, which will affect their preventive behaviour. This can greatly benefit transport companies, as the results show that the vehicles are utilised more efficiently and in a better technical condition. The results therefore represent an important guideline for improving maintenance in transport companies, developing new competencies of truck drivers and upgrading existing knowledge in the development of modern operation maintenance.

ŠKERLIČ S, SOKOLOVSKIJ E, ERČULJ V. **Obsługa techniczna samochodów ciężarowych: międzynarodowe badanie kierowców ciężarówek**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 493–500, http://dx.doi.org/10.17531/ein.2020.3.12.

Od czasu wdrożenia nowoczesnych metod obsługi technicznej, kierowcy powinni zostać zintegrowani z całym systemem obsługi technicznej, aby przejąć kompetencje zawodowe pracowników obsługi technicznej. W tym celu przeprowadzono międzynarodowe badanie na próbie 249 kierowców ciężarówek w celu ustalenia, w jaki sposób obsługa techniczna w firmach transportowych wpływa na rolę kierowców ciężarówek w utrzymaniu floty. Na podstawie opracowanego modelu równania strukturalnego (SEM) ustalono, że zbudowanie efektywnej infrastruktury serwisowej w firmach transportowych umożliwia aktywny udział kierowców ciężarówek w systemie obsługi technicznej firmy. Kierowcy są zdolne do naprawy drobnych usterek, ale tylko wtedy, gdy są proaktywni, co wpłynie na ich zachowanie zapobiegawcze. Może to być bardzo korzystne dla firm transportowych, ponieważ wyniki pokazują, że pojazdy są użytkowane bardziej wydajnie i w lepszym stanie technicznym. Wyniki stanowią zatem ważną wskazówkę dotyczącą poprawy obsługi technicznej w firmach transportowych poprzez rozwój nowych kompetencji kierowców ciężarówek i poszerzanie istniejącej wiedzy w zakresie rozwoju nowoczesnej obsługi technicznej.

PASZKOWSKI W. **Modeling of vibroacoustic phenomena using the method of parameterizing the audio signal**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 501–507, http://dx.doi.org/10.17531/ein.2020.3.13.

The article proposes an original way of modeling vibroacoustic phenomena of exploited machines/devices using the method of audio parameterization. This method extends the current approach to this type of research and consists in taking into account the psychoacoustic effects associated with the emission of vibroacoustic energy. The proposed solution is based on the determination of mel-cepstral coefficients of the examined signal and its classification, due to the impact of noise. It was presented verification of the method on the example of studies on the impact of road noise sources.

PASZKOWSKI W. **Modelowanie zjawisk wibroakustycznych z zastosowaniem metody parametryzacji sygnału fonicznego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 501–507, http://dx.doi.org/10.17531/ein.2020.3.13.

W artykule zaproponowano oryginalny sposób modelowania zjawisk wibroakustycznych eksploatowanych maszyn/urządzeń z zastosowaniem metody parametryzacji sygnału fonicznego. Sposób ten rozszerza dotychczasowe podejście do tego rodzaju badań i polega na uwzględnianiu efektów psychoakustycznych towarzyszących emisji energii wibroakustycznej. Proponowane rozwiązanie opiera się na wyznaczeniu współczynników mel-cepstralnych badanego sygnału i jego klasyfikacji, ze względu na oddziaływanie hałasu. Przedstawiono weryfikację zastosowania metody na przykładzie badań oddziaływania źródeł hałasu drogowego.

LIANG H, MI J, BAI L, CHENG Y. **Imprecise sensitivity analysis of system reliability based on the Bayesian network and probability box**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 508–519, http://dx.doi.org/10.17531/ein.2020.3.14.

Sensitivity analysis measures how changes in system inputs affect outputs. Previously, a large amount of sensitivity analysis research was relevant to the precise probability that is regarded as an ideal condition of engineering. Due to insufficient test samples and the low accuracy of test data, system reliability with hybrid uncertainty is difficult to be described as a precise value. As a profusion of highly integrated electromechanical equipment is applied in modern life, it is impossible to apply sufficient resources to eliminate the stochastic property of every component, which necessitates

LIANG H, MI J, BAI L, CHENG Y. **Niedokładna analiza czułościowa niezawodności systemu w oparciu o sieć bayesowską i pole prawdopodobieństwa (p-box)**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 508–519, http://dx.doi.org/10.17531/ein.2020.3.14.

Celem analizy czułościowej jest badanie w jakim stopniu zmiany danych wejściowych systemu wpływają na dane wyjściowe. Dotychczasowe badania z wykorzystaniem analizy czułościowej były związane z dokładnym prawdopodobieństwem postrzeganym w inżynierii jako warunek idealny. Przy niewystarczającej wielkości badanej próby i niskiej dokładności danych testowych, niezawodność systemu o hybrydowej niepewności trudno opisać w sposób dokładny. Biorąc pod uwagę fakt, że we współczesnym świecie wykorzystuje się duże ilości wysoce zintegrowanych urządzeń elektromechanicznych,

the identification of highly sensitive components to efficiently reduce imprecision. Hence, based on the theory of imprecise probability, imprecise sensitivity analysis has become a popular research topic in the last decade. In this paper, a method for uncertain system reliability and imprecise sensitivity analysis is proposed based on a Bayesian network, a probability box and the pinching method. The feasibility and accuracy of the combined method are fully verified through the evaluation and analysis of a numerical example and a case study of an electromechanical system, and the highly sensitive components that heavily influence the imprecision of system outputs are accurately identified.

niemożliwa jest alokacja wystarczających zasobów w celu wyeliminowania właściwości stochastycznych każdego elementu. Oznacza to, że aby zredukować niedokładność, konieczna jest identyfikacja komponentów o wysokiej czułości. Dlatego też popularnym przedmiotem badań ostatniej dekady stała się niedokładna analiza czułości, bazująca na teorii niedokładnego prawdopodobieństwa. W artykule zaproponowano metodę analizy niezawodności niepewnego systemu jak również niedokładnej analizy czułościowej w oparciu o sieć bayesowską, pole prawdopodobieństwa i metodę pinch point. Możliwość wykorzystania i dokładność metody zostały w pełni potwierdzone na podstawie przykładu liczbowego jak również studium przypadku systemu elektromechanicznego; proponowana metoda pozwoliła na poprawne określenie wysoce czułych elementów systemu, które w dużym stopniu wpływają na niedokładność danych wyjściowych układu.

MA B, YAN S, WANG X, CHEN J, ZHENG C. **Similarity-based failure threshold determination for system residual life prediction**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 520–529, http://dx.doi.org/10.17531/ein.2020.3.15.

An accurate determination of the system failure threshold is an essential requirement in achieving an appropriate system residual life prediction and a reasonable planned maintenance strategy optimization afterward for degradation systems. This paper proposes a failure threshold determination method based on quantitative measurement of the similarity between the operating system and the historical systems. The similarity is formulated by a weighted average function and then calculated by a convex quadratic formulation to minimizing the variance between the operating system and the historical systems. With an accurate determination of the system failure threshold in real-time, a better prediction of the residual life for the operating system is achieved. Finally, a real case study for several power-shift steering transmission systems monitored using oil spectral analysis is adopted to illustrate and numerically compare the improved performance of the proposed method.

MA B, YAN S, WANG X, CHEN J, ZHENG C. **Określanie progu awarii na podstawie podobieństwa jako metoda pozwalająca na przewidywanie trwałości resztkowej systemu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 520–529, http://dx.doi.org/10.17531/ein.2020.3.15.

W przypadku systemów podlegających degradacji, dokładne określenie progu awarii systemu stanowi niezbędny warunek dokonania trafnej prognozy jego trwałości resztkowej oraz późniejszej optymalizacji strategii konserwacji rutynowych. W artykule zaproponowano metodę wyznaczania progu awarii opartą na ilościowym pomiarze podobieństwa między systemem użytkowanym obecnie a systemami użytkowanymi uprzednio. Podobieństwo formułuje się na podstawie funkcji średniej ważonej, a następnie oblicza na podstawie wypukłej formy kwadratowej w celu zminimalizowania wariancji między obecnie użytkowanym systemem a uprzednimi systemami. Dzięki dokładnemu określeniu progu awarii systemu w czasie rzeczywistym uzyskuje się lepszą prognostykę trwałości resztkowej obecnie użytkowanego systemu. W końcowej części pracy, w celu zilustrowania i numerycznego porównania ulepszonej wydajności proponowanej metody, zaprezentowano studium przypadku obejmujące kilka układów przeniesienia napędu monitorowanych przy użyciu analizy spektralnej oleju.

CERAN B, ORŁOWSKA A, KROCHMALNY K. **The method of determining PEMFC fuel cell stack performance decrease rate based on the voltage-current characteristic shift**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 530–535, http://dx.doi.org/10.17531/ein.2020.3.16.

The article presents mathematical model designed to estimate the rate of performance decrease in fuel cell stack. The fuel cell stack performance decrease rate is determined on the basis of stack average voltage measurements. The proposed model is used to determine power curve as well as exploitation indicators of fuel cell stack with a nominal power of 50 kW after 14 000 hours of continuous operation. The model is also used to determine the average voltage drop for the eleven-year fuel cell stack with a nominal power of 1,2 kW. In both studies, the values of exploitation indicators as well as their differences in relation to nominal values are determined.

CERAN B, ORŁOWSKA A, KROCHMALNY K. **Metoda wyznaczania szybkości spadku wydajności stosu ogniw paliwowych typu PEMFC na podstawie przesunięcia charakterystyki napięciowo-prądowej**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 530–535, http://dx.doi.org/10.17531/ein.2020.3.16.

Artykuł przedstawia model matematyczny przeznaczony do wyznaczenia szybkości spadku wydajności stosu ogniw paliwowych. Szybkość spadku wydajności stosu ogniw jest wyznaczana na podstawie wartości napięcia średniego stosu. Zaproponowany model wykorzystano do wyznaczenia krzywej mocy i wskaźników eksploatacyjnych stosu ogniw paliwowych o mocy nominalnej 50 kW po 14 000 h ciągłej pracy. Model wykorzystano także do wyznaczenia szybkości zmiany wartości napięcia średniego jedenastoletniego stosu ogniw paliwowych o mocy 1,2 kW. W obu badaniach wyznaczono wartości wskaźników eksploatacyjnych oraz ich różnice względem wielkości nominalnych.

WALKER P, DOROSZUK B, KRÓL R. **Analysis of ore flow through longitudinal belt conveyor transfer point**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 536–543, http://dx.doi.org/10.17531/ein.2020.3.17.

A transfer point is an element of a belt conveyor prone to increased energy losses and to the risk of failure. It is also a location in which the receiving belt is particularly susceptible to damage. Except failure-free operation, a transfer point should offer minimal belt resistances to motion by ensuring that the transported material is placed centrally on the receiving belt, both spillage of the material and blockages are prevented, the process of particle defragmentation is limited, and also that noise and dust emissions to the environment are reduced. Ensuring that the above requirements are met requires inter alia the use of advanced simulation tests. The article analyzes the flow of ore particles stream through a longitudinal transfer point used in an underground copper ore mine. Discrete Element Method was used to identify the phenomena which occur while transferring ore onto the receiving conveyor. The research allowed key variables affecting the transfer point performance to be identified. It also resulted in a proposal of actions which can improve the performance of the transfer point and which are focused on saving energy and on minimizing the damage and wear of the receiving belt.

WALKER P, DOROSZUK B, KRÓL R. **Badania symulacyjne przesypu wzdłużnego przenośnika taśmowego**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 536–543, http://dx.doi.org/10.17531/ein.2020.3.17.

Przesyp jest miejscem przenośnika taśmowego, w którym pojawia się ryzyko wystąpienia awarii, występują straty energii oraz może dochodzić do uszkodzenia taśmy odbierającej urobek. Poza bezawaryjnym funkcjonowaniem przesyp powinien dla zminimalizowania oporów ruchu taśmy zapewnić także centralne podawanie urobku, zapobiegać rozsypywaniu się transportowanego materiału, nie dopuszczać do powstawania zatorów, ograniczać proces defragmentacji urobku, a także minimalizować emisję hałasu oraz pyłów do otoczenia. Zapewnienie stawianych wymagań wiąże się z koniecznością stosowania m.in. zaawansowanych badań symulacyjnych. W artykule przeprowadzono analizę przepływu strugi urobku przez wybrany przesyp wzdłużny, stosowany w podziemnej kopalni rud miedzi. Przy użyciu metody elementów dyskretnych DEM dokonano oceny zjawisk zachodzących podczas przesypywania rudy na przenośnik odbierający, wskazano kluczowe zmienne opisujące jego pracę, a także zaproponowano działania udoskonalające pracę przesypu, zorientowane na zwiększenia jego energooszczędności oraz zmniejszenia negatywnego oddziaływania transportowanego nosiwa na taśmę przenośnika odbierającego.

KARABACAK YE, GÜRSEL ÖZMEN N, GÜMÜŞEL L. **Worm gear condition monitoring and fault detection from thermal images via deep learning method**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 544–556, http://dx.doi.org/10.17531/ein.2020.3.18.

Worm gearboxes (WG) are often preferred, because of their high torque, quickly reducing speed capacity and good meshing effectiveness, in many industrial applications. However, WGs may face with some serious problems like high temperature at the speed reducer, gear wearing, pitting, scoring, fractures and damages. In order to prevent any damage, loss of time and money, it is an important issue to detect and classify the faults of WGs and develop the maintenance plans accordingly. The present study addresses the application of the deep learning method, convolutional neural network (CNN), in the field of thermal imaging that were gathered from a test rig

KARABACAK YE, GÜRSEL ÖZMEN N, GÜMÜŞEL L. **Monitorowanie stanu i wykrywanie błędów przekładni ślimakowej na podstawie termogramów z wykorzystaniem metody głębokiego uczenia**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 544–556, http://dx.doi.org/10.17531/ein.2020.3.18.

W wielu zastosowaniach przemysłowych preferuje się przekładnie ślimakowe, ze względu na ich wysoki moment obrotowy, możliwość szybkiej redukcji prędkości i dobrą sprawność zazębienia. Jednakże przekładnie tego typu narażone są często na poważne problemy, takie jak wysoka temperatura przy reduktorze prędkości czy też zużycie, pitting (wżery), zatarcie, pęknięcie lub uszkodzenie kół zębatych. Zapobiec takim uszkodzeniom, i związanym z nimi stratom finansowym i czasowym, można poprzez wykrywanie i klasyfikowanie błędów przekładni i odpowiednie opracowanie planów konserwacji. Niniejsze badanie

operating on different loads and speeds. Deep learning approaches, have proven their powerful capability to exploit faulty information from big data and make intelligently diagnostic decisions. Studies concerning the condition monitoring of WGs in the literature are limited. This is the first study on WGs with infrared thermography rather than vibration and sound measurements which have some deficiencies about hardware requirements, restricted measurement abilities and noisy signals. For comparison, CNN was also trained, with vibration and sound data which were collected from the healthy and faulty WGs. The results of fault diagnosis show that thermal image based CNN model on WG has achieved 100% success rate whereas the vibration performance was 83.3 % and sound performance was 81.7%. As a result, thermal image based CNN model showed a better diagnosing performance than the others for WGs. Moreover, condition monitoring of WGs, can be performed correctly with less measurement costs via thermal imaging methods.

dotyczy zastosowania metody głębokiego uczenia oraz splotowych sieci neuronowych (SSN) do monitoringu stanu przekładni na podstawie termogramów zarejestrowanych na stanowisku testowym pracującym przy różnych obciążeniach i prędkościach. Podejścia oparte na uczeniu głębokim umożliwiają efektywne wykorzystanie informacji o błędach pochodzących z dużych zbiorów danych i podejmowanie trafnych decyzji diagnostycznych. Niewiele z dostępnych publikacji poświęconych jest monitorowaniu stanu przekładni ślimakowych. Niniejsza praca jako pierwsza przedstawia badania przekładni ślimakowej z zastosowaniem termografii zamiast zwyczajowo prowadzonych pomiarów drgań i dźwięku, które mają pewne wady dotyczące wymagań sprzętowych, ograniczonych możliwości pomiarowych i głośności sygnałów. SNN opartą na danych termicznych porównano z siecią, którą uczono na zbiorach danych wibracyjnych i akustycznych pochodzących z prawidłowo działających i uszkodzonych przekładni ślimakowych. Wyniki diagnostyki uszkodzeń pokazują, że model SSN przekładni ślimakowej oparty na obrazie termicznym osiągnął stuprocentową (100%) skuteczność, podczas gdy skuteczność modeli opartych na danych wibracyjnych i akustycznych wyniosła, odpowiednio, 83,3% i 81,7%. Tym samym, model SNN oparty na obrazie termicznym pozwalał na trafniejsze diagnozowanie przekładni ślimakowej niż pozostałe modele. Ponadto zastosowanie metod opartych na termografii pozwala na poprawne monitorowanie stanu przy niższych kosztach pomiaru.

UŁANOWICZ L, JASTRZĘBSKI G, SZCZEPANIAK P. **Method for estimating the durability of aviation hydraulic drives**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 557–564, http://dx.doi.org/10.17531/ein.2020.3.19.

Throughout previous practice, estimating the life of aviation hydraulic drive assemblies has been utilizing a variant, which requires conducting long-lasting studies of the drive assemblies until they move to the unfitness state. Such studies, which enable estimating life a posteriori, are costly and long-lasting. Hence the need to look for new strategies for estimating life. The article presents a method of estimating the durability of a hydraulic drive assembly based on the control of its change in technical condition. Inspection of the technical condition enables timely detection of the condition before the emergency hydraulic assembly. The novelty of the method is to use, to detect the condition before the emergency team, the principle of determining the pre-emptive control parameter tolerance. Pre-emptive tolerances are a set of control parameter values between threshold levels and pre-emergency (allowable) levels. The intensity of depletion of durability (intensity of aging, wear) is random. The paper presents a stochastic description of the control parameter change and the resulting empirical relationships between the control parameter verification time probability density (verification periodicity) and the control parameter value change probability density. The inter-relations between these two functions were described. It also presents empirical relationships enabling the determination of the permissible value for the control parameters and the periodicity of the control parameter checks after exceeding the limit value. An example of estimating the life of a hydraulic piston pump on-board an aircraft operated in the Polish Air Forces was shown. The permissible values and the time for the first control parameter verification after exceeding the limit value were determined for selected control parameters of the hydraulic pump. The proposed method binds life (fitness time) with the physical wear mechanisms concerning the assemblies. It can be applied in work aimed at determining the resource life of technical equipment. Furthermore, it enables utilizing technical equipment according to a technical state strategy with monitoring the parameters.

UŁANOWICZ L, JASTRZĘBSKI G, SZCZEPANIAK P. **Metoda szacowania trwałości lotniczych napędów hydraulicznych**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 557–564, http://dx.doi.org/10.17531/ein.2020.3.19.

W dotychczasowej praktyce szacowania trwałości zespołów lotniczych napędów hydraulicznych stosowany jest wariant, który wymaga prowadzenia długotrwałych badań zespołów napędu do czasu ich przejścia w stan niezdatności. Badania tego typu, umożliwiające szacowanie trwałości a posteriori, są kosztowne i długotrwałe. Istnieje więc potrzeba poszukiwania nowych strategii szacowania trwałości. W artykule zaprezentowano metodę szacowania trwałości zespołu napędu hydraulicznego opartą o kontrolę jego zmiany stanu technicznego. Kontrola stanu technicznego umożliwia wykrycie we właściwym czasie stanu przed awaryjnego zespołu hydraulicznego. Novum metody jest wykorzystanie, do wykrycie stanu przed awaryjnego zespołu, zasady wyznaczania uprzedzających tolerancji parametru kontrolnego. Tolerancje uprzedzające stanowią zbiór wartości parametru kontrolnego zawartych między poziomami granicznym i przed awaryjnym (dopuszczalnym). Intensywność wyczerpywania się trwałości (intensywności starzenia, zużywania) ma losowy charakter. W artykule przedstawiono stochastyczny opis zmiany parametru kontrolnego oraz wynikające z niego empiryczne zależności funkcji gęstości prawdopodobieństwa czasu przeprowadzania sprawdzeń parametru kontrolnego (okresowość kontroli) i funkcji gęstości prawdopodobieństwa zmiany wartości parametru kontrolnego. Opisano wzajemne związki obu tych funkcji. Przedstawiono zależności umożliwiające wyznaczenie wartości dopuszczalnej parametru kontrolnego i okresowość sprawdzeń parametru kontrolnego po przekroczeniu wartości dopuszczalnej. Zaprezentowano przykład szacowania trwałości tłoczkowej pompy hydraulicznej z samolotu użytkowanego w Siłach Zbrojnych RP. Dla wybranych parametrów kontrolnych pompy hydraulicznej wyznaczono ich wartości dopuszczalne oraz czas pierwszej kontroli parametru kontrolnego po przekroczeniu wartości dopuszczalnej. Zaprezentowana metoda wiąże trwałość z fizycznymi mechanizmami zużywania się zespołów. Przedstawiona metoda może być wykorzystana w pracach mających na celu określanie zasobu pracy urządzeń technicznych. Umożliwia ona użytkowanie urządzeń technicznych według strategii stanu technicznego z kontrolowaniem parametrów.

KOTELKO M, MACDONALD M, KULATUNGAMP, MARSZALEK Z. **Upper-bound estimation of load-carrying capacity of perforated cold-formed thin-walled steel lipped channel columns under compression loading**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 565–573, http://dx.doi.org/10.17531/ein.2020.3.20.

Upper-bound estimation of the load-capacity of cold-formed steel sections (TWCFS) with perforations, subjected to axial compression is presented. The estimation is performed on the basis of the Yield Line Analysis ((YLA). TWCFS lipped channel sections with two sets of perforations (on the web and on the flanges) are under investigation. The comparison of experimental results, FE simulation results, European code ultimate strength predictions and upper-bound estimation based on YLA approach is carried out and presented. Some conclusions concerning an applicability of the YLA approach for ultimate strength prediction of perforated TWCFS structural members are derived.

KOTELKO M, MACDONALD M, KULATUNGAMP, MARSZALEK Z. **Oszacowanie górne nośności perforowanych zimno formowanych prętów cienkościennych poddanych ściskaniu**. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 565–573, http://dx.doi.org/10.17531/ein.2020.3.20.

W artykule przedstawiono wyniki górnego oszacowania nośności cienkościennych prętów zimno formowanych z perforacjami, poddanych osiowemu ściskaniu. Oszacowanie to jest oparte na metodzie załomów plastycznych. Rozpatrywano dwa warianty perforacji (środnika i pasów) cienkościennych prętów ceowych z żebrami końcowymi. Przeprowadzono analizę porównawczą wyników eksperymentu, wyników symulacji numerycznych MES oraz wyników obliczeń wg wzorów normatywnych normy europejskiej z wynikami oszacowania górnego nośności opartego na metodzie załomów plastycznych. Sformułowano wnioski dotyczące możliwości zastosowania tej metody do szacowania nośności cienkościennych prętów zimno formowanych z perforacjami.

SCIENCE AND TECHNOLOGY

Article citation info:
LIU D, WANG S, TOMOVIC MM. Degradation modeling method for rotary lip seal based on failure mechanism analysis and stochastic process. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 381–390, http://dx.doi.org/10.17531/ein.2020.3.1.

Di LIU
Shaoping WANG
Mileta M. TOMOVIC

DEGRADATION MODELING METHOD FOR ROTARY LIP SEAL BASED ON FAILURE MECHANISM ANALYSIS AND STOCHASTIC PROCESS

METODA MODELOWANIA DEGRADACJI OBROTOWEGO USZCZELNIENIA WARGOWEGO W OPARCIU O ANALIZĘ MECHANIZMU USZKODZENIA I PROCES STOCHASTYCZNY

Rotary lip seal is widely used in aircraft and its performance affects the safety of the aircraft. Hence, it is necessary to estimate useful lifetime and reliability of the seal. Degradation of rotary lip seal is always with random effects, which cannot be considered by theoretical failure mechanism analysis. Hence, in order to consider the random effects of rotary lip seal degradation, stochastic processes are applied. Furthermore, considering the monotonic degradation of the seal, Gamma process and inverse Gaussian process are selected as the candidate processes. To combine the candidate processes, Bayesian model averaging is introduced. Based on the failure mechanism analysis and numerical simulation, the theoretical wear path is predicted and corresponding linearization method is proposed. The measured degradation data is converted and the seal wear process is transformed to a linear degradation process. The model parameters and model probabilities are evaluated by fully Bayesian inference method. The effectiveness of the proposed method is verified by comparing the predicting degradation and experimental observations. The proposed method can be used to evaluate reliability and useful lifetime of rotary lip seal. According to sensitivity analysis, an effective way to improve lifetime and reliability of the seal is to increase the wear depth threshold.

Keywords: *Rotary lip seal, Failure mechanism analysis, Stochastic process, Degradation modeling, Bayesian model averaging.*

Obrotowe uszczelnienia wargowe znajdują szerokie zastosowanie w samolotach, a ich sprawność wpływa na bezpieczeństwo statków powietrznych. Oznacza to, iż szacowanie żywotności i niezawodności tego rodzaju uszczelnień ma kluczowe znaczenie. Degradacja obrotowego uszczelnienia wargowego jest zawsze związana z efektami losowymi, których nie uwzględnia teoretyczna analiza mechanizmu uszkodzenia. Dlatego też do oceny efektów losowych degradacji obrotowego uszczelnienia wargowego wykorzystuje się procesy stochastyczne, takie jak proces Gamma czy odwrotny proces Gaussa. W przedstawionej pracy, wybrane procesy degradacji łączono za pomocą metody bayesowskiego uśredniania modeli. Na podstawie analizy mechanizmów uszkodzeń i symulacji numerycznej, konwertowano uzyskane w pomiarach dane degradacyjne, co pozwoliło na przekształcenie procesu degradacji obrotowego uszczelnienia wargowego w proces liniowy. Parametry modelu i prawdopodobieństwa oceniano za pomocą metody pełnego wnioskowania bayesowskiego na podstawie obserwacji degradacji. Skuteczność przedstawionej metody weryfikowano porównując przewidywane i obserwowane wartości degradacji. Proponowaną metodę można wykorzystywać do oceny niezawodności i żywotności obrotowego uszczelnienia wargowego. Przeprowadzona analiza czułości pokazuje, że skutecznym sposobem na poprawę żywotności i niezawodności omawianego typu uszczelnienia jest zwiększenie progu uszkodzenia w postaci maksymalnej głębokości zużycia.

Słowa kluczowe: *Obrotowe uszczelnienie wargowe, mechanizm uszkodzenia, proces stochastyczny, modelowanie degradacji, bayesowskie uśrednianie modeli.*

1. Introduction

Dynamic seals are commonly applied in aircrafts, such as hydraulic system and fuel system [20, 37]. In order to meet the requirements of modern air-craft fuel system and hydraulic system, such as high speed and high sealed pressure, dynamic seals have been paid more and more attention [9, 14, 21]. Furthermore, rotary lip seal is always used in fuel system, resulting in the failure of the rotary lip seal may lead to an incredible disaster. Hence, rotary lip seal is focused in the presented research.

A fundamental sealing mechanism of rotary lip seal with reverse pumping action is based on the profile of sealing lip and the shaft-seal interference [15]. By pumping fluid from the air-side to the liquid-side, this reverse pumping action can prevent leakage [5, 29, 30, 31]. Reverse pumping rate is one of the most important sealing performance indicators. Leak is likely to occur and the seal is likely to fail when reverse pumping rate is small enough. Furthermore, without the shaft-seal interference, the reverse pumping action will be destroyed and the seal will fail.

Several researches on analyzing degradation mechanism of the rotary lip seals have been published. Guo et al. have performed a series of experiments to analyze degradation of the seal during storage, aging in oil and using in system [6,7,8]. The authors argue that the rubber aging is the most important failure factor for the seals during storage. The roughness of sealing lip surface deceases with rubber aging, resulting in the reverse pumping action weakening. Sealing lip wear is the most important failure factor for rotary lip seal during using in system. Seal wear changes the lip profile, decreases the shaft-seal interference and weakens the reverse pumping action. A storage lifetime assessment method has been proposed in our previously published paper [10]. Hence, the useful lifetime of rotary lip seal and sealing lip wear are focused in the presented research.

Based on finite elements (FE) analysis method and Archard wear equation, several researches have been published to numerically study the seal wear process. To analyze the wear of the seal used in downhole equipment, Li et al. [35] presented a numerical analysis method for hydraulic seals. The high temperature and high pressure conditions are focus. The proposed method has been validated by comparing the simulation results with experimental observations. Based on the above numerical method, Li et al. [36] proposed a thermal-structural coupled FE analysis method. The thermal behavior is coupled with structural analysis. A mesh reconstruction strategy is used to describe the evolution of seal geometry caused by wear. The proposed method can be used to provide a suggestion for the applications considering the temperature effects. Considering the effects of lubricating characteristics on the seal wear, we have proposed a multiscale wear simulation method for rotary lip seal [11]. The proposed method has been verified by experimental observations and used to study the relationship between shaft roughness and seal wear. Based on the above method, the relationship between the shaft texture and seal wear has been numerically studied in [12]. This work provides insight into the relationship between geometric features of textured shaft and the seal wear. Based on the above wear simulation method, the theoretical degradation path of rotary lip seal can be modeled and the theoretical lifetime of the seal can also be predicted based on the simulation results [3, 4].

Based on failure mechanism analysis, the theoretical wear path and theoretical lifetime can be predicted. However, the degradation of rotary lip seal is always with random effects, which cannot be considered in theoretical failure mechanism analysis. Stochastic processes are normally introduced in degradation modeling to take into account the random effects [38]. Hence, in order to consider random effects of rotary lip seal degradation, stochastic processes are applied in the present research [13, 23].

Using stochastic process to model the degradation of rotary lip seal, the degradation process, degradation path and model parameters need to determine. The degradation process of rotary lip seal is a monotonous process. Generally, inverse Gaussian (IG) process and Gamma process are suitable to describe monotonous processes [26, 28]. Hence, IG process and Gamma process are used to model the degradation process of rotary lip seal in the presented research. Generally, under enough sample conditions, Akaike information criterion (AIC) and Bayesian information criterion (BIC) can be used to select the best degradation process among the candidate processes based on degradation dataset [16, 17]. However, due to the small sample condition of rotary lip seal, it is difficult to select the best degradation process based on AIC or BIC. Normally, Bayesian model averaging (BMA) method can be used to combine the candidate degradation processes [18, 27]. Hence, BMA method is used to combine Gamma process and IG process in the presented research.

The best degradation path among the candidate paths also can be selected by AIC and BIC under enough sample conditions. For example, in [24], the best degradation path among the candidate degradation paths, including paths with constant, monotonic, and S-shaped degradation rates, is selected based on AIC. However, similar to degradation process, under small sample conditions, it is also difficult to determine degradation path based on degradation dataset. Generally, degradation path reflects the theoretical degradation feature and can be predicted based on failure mechanism analysis [33]. Hence, in the presented research, the theoretical degradation path of rotary lip seal is predicted based on failure analysis and numerical simulation. Furthermore, the degradation process is always transformed to a linear degradation process by converting the degradation observations based on a linearization approach, which is given based on the above theoretical path [1, 34]. Generally, the model parameters and process probabilities can be evaluated by Bayesian inference method with Monte Carlo simulation method [2, 19, 25].

As discussed above, based on failure mechanism analysis and stochastic process, a degradation modeling method for rotary lip seals are proposed. Comparing to the previously published researches on degradation modeling and reliability analysis on rotary lip seal, the distinguished features of this paper are as follows: (1) Stochastic process is introduced to consider the randomness of the seal degradation. (2) Degradation path is predicted based on the failure mechanism analysis and numerical simulation. Based on the calculated theoretical path, the measured degradation data is converted and the rotary lip seal degradation process is transformed to a linear process. (3) BMA method is applied to handle process uncertainty issue, evaluate the model parameters and process probabilities.

Furthermore, the presented research is further study of the authors' previously published paper [13] and corresponding application on rotary lip seal. To meet engineering practice, in the presented research, the method to predict the degradation path and corresponding linearization method are proposed for rotary lip seal. The degradation modeling and reliability analysis method is not be limited to linear degradation process anymore compared to the previously published paper. This improvement makes the stochastic process based degradation modeling method become more practicable and available in engineering practice.

The rest of this paper is as follows: Section 2 presents the stochastic process based degradation modeling method. In Section 3, the theoretical degradation path and the linearization method are proposed based on failure mechanism analysis. In Section 4, the model parameters and process probabilities are estimated by Bayesian inference method. In Section 5, the effectiveness of the proposed method is demonstrated. In Section 6, discussion and conclusions are presented.

2. Stochastic process based degradation modeling method

2.1. Bases of stochastic process

A degradation process, which can be described as stochastic process, has the following properties:

(a) $Y(t)$ has independent increments, e.g. $Y(t_4)$ - $Y(t_3)$ and $Y(t_2)$ - $Y(t_1)$ are independent only if $t_4 > t_3 > t_2 > t_1$;

(b) The initial degradation $Y(0)\equiv 0$;

(c) The increments ΔY follow

$$\Delta Y(t) \sim f\left(\Delta\Lambda(t), \theta\right) \tag{1}$$

where $\Delta Y(t)=Y(t+\Delta t)-Y(t)$ and $\Delta\Lambda(t)=\Lambda(t+\Delta t)-\Lambda(t)$, $\Lambda(t)$ is a monotonic function, and $\boldsymbol{\theta}$ is model parameter vector.

The form of probability density function (PDF), f, is determined by the selected degradation process, Gamma process and IG process in the presented research. The form of monotonic increasing function $\Lambda(t)$ is determined by the degradation path. In the presented research, it is transformed to linear process based on linearization method and the theoretical degradation path, which is calculated by degradation mechanism analysis and numerical simulation. The model parameters and model probabilities can be evaluated based on degradation dataset, including degradation observations vector $\mathbf{Y}$ and observation time vector $\mathbf{T}$:

$$\mathbf{Y} = \left[y_1, y_2, \ldots y_n\right] \tag{2}$$

$$\mathbf{T} = \left[t_1, t_2, \ldots t_n\right] \tag{3}$$

where y_i means i-th degradation observation, t_i means i-th corresponding observation time and n is the observation size.

2.2. Linear IG process based degradation model

The degradation process of the seal is transformed to linear process. In linear IG process, the degradation Y follows IG distribution, as:

$$Y(t) \sim \mathrm{IG}\left(\mu_{\mathrm{IG}} t, \lambda_{\mathrm{IG}} t^2\right) \tag{4}$$

where μ_{IG} and λ_{IG} are the process parameters. The PDF of IG distribution, x~IG(a_{IG},b_{IG}), is given by:

$$f_{\mathrm{IG}}\left(x \middle| a_{\mathrm{IG}}, \mathrm{b}_{\mathrm{IG}}\right) = \sqrt{\frac{b_{\mathrm{IG}}}{2\pi x^3}} \exp\left[-\frac{b\left(x-a_{\mathrm{IG}}\right)^2}{2{a_{\mathrm{IG}}}^2 x}\right], x>0 \tag{5}$$

Hence, the corresponding PDF of the degradation y is given by:

$$f_{\mathrm{IG}}\left(y \middle| \mu_{\mathrm{IG}} t, \lambda_{\mathrm{IG}} t^2\right) = \sqrt{\frac{\lambda_{\mathrm{IG}} t^2}{2\pi y^3}} \exp\left[-\frac{\lambda_{\mathrm{IG}}\left(y-\mu_{\mathrm{IG}} t\right)^2}{2\mu_{\mathrm{IG}}^2 y}\right] \tag{6}$$

Given the predefined failure threshold, D, the corresponding reliability function is given by:

$$R_{\mathrm{IG}}\left(t \middle| \mu_{\mathrm{IG}} t, \lambda_{\mathrm{IG}} t^2\right) = \Phi\left[\sqrt{\frac{\lambda_{\mathrm{IG}}}{D}}\left(\frac{D}{\mu_{\mathrm{IG}}}-t\right)\right] + \exp\left(\frac{2\lambda_{\mathrm{IG}} t}{\mu_{\mathrm{IG}}}\right)\Phi\left[-\sqrt{\frac{\lambda_{\mathrm{IG}}}{D}}\left(\frac{D}{\mu_{\mathrm{IG}}}+t\right)\right] \tag{7}$$

where Φ is cumulative probability function (CDF) of standard Gaussian distribution.

The corresponding mean time to failure (MTTF) is given by [22]:

$$MTTF_{\mathrm{IG}}\left(\theta_{\mathrm{IG}}\right) = \left(\frac{D}{\mu_{\mathrm{IG}}} + \frac{\mu_{\mathrm{IG}}}{\lambda_{\mathrm{IG}}}\right)\Phi\left(\frac{\sqrt{\lambda_{\mathrm{IG}} D}}{\mu_{\mathrm{IG}}}\right) + \sqrt{\frac{D}{\lambda_{\mathrm{IG}}}}\phi\left(\frac{\sqrt{\lambda_{\mathrm{IG}} D}}{\mu_{\mathrm{IG}}}\right) - \frac{\mu_{\mathrm{IG}}}{2\lambda_{\mathrm{IG}}} \tag{8}$$

where $\phi(.)$ is PDF of standard Gaussian distribution.

2.3. Linear Gamma process based degradation model

The degradation process of the seal is transformed to linear process. In linear Gamma process, the degradation Y follows Gamma distribution, as:

$$Y(t) \sim \mathrm{Ga}\left(\mu_{\mathrm{Ga}} t, \lambda_{\mathrm{Ga}}\right) \tag{9}$$

where μ_{Ga} and λ_{Ga} are the process parameters. The PDF of Gamma distribution, y~Ga(a_{Ga},b_{Ga}), is given by:

$$f_{\mathrm{Ga}}\left(y \middle| a_{\mathrm{Ga}}, \mathrm{b}_{\mathrm{Ga}}\right) = \frac{1}{\Gamma\left(a_{\mathrm{Ga}}\right){b_{\mathrm{Ga}}}^{a_{\mathrm{Ga}}}} y^{a_{\mathrm{Ga}}-1} \exp\left(-y / b_{\mathrm{Ga}}\right) \tag{10}$$

where $\Gamma(\cdot)$ is Gamma function.

Hence, the corresponding PDF of the degradation is given by:

$$f_{\mathrm{Ga}}\left(y \middle| \mu_{\mathrm{Ga}} t, \lambda_{\mathrm{Ga}}\right) = \frac{y^{\mu_{\mathrm{Ga}} t-1}}{\Gamma\left(\mu_{\mathrm{Ga}} t\right){\lambda_{\mathrm{Ga}}}^{\mu_{\mathrm{Ga}} t}} \exp\left(-\frac{y}{\lambda_{\mathrm{Ga}}}\right) \tag{11}$$

The corresponding reliability function is given by:

$$R_{\mathrm{Ga}}\left(t \middle| \mu_{\mathrm{Ga}} t, \lambda_{\mathrm{Ga}}\right) = 1 - \frac{\Gamma\left(\mu_{\mathrm{Ga}} t, D/\lambda_{\mathrm{Ga}}\right)}{\Gamma\left(\mu_{\mathrm{Ga}} t\right)} \tag{12}$$

The corresponding MTTF is given by [32]:

$$MTTF_{\mathrm{Ga}}\left(\theta_{\mathrm{Ga}}\right) = \frac{D}{\mu_{\mathrm{Ga}}\lambda_{\mathrm{Ga}}} + \frac{1}{2\mu_{\mathrm{Ga}}} \tag{13}$$

2.4. BMA model fusion method

Considering monotonic degradation of rotary lip seal, IG process and Gamma process are selected as candidate degradation processes. As discussed above, under enough sample conditions, the best fitting degradation process among the candidate processes can be selected by AIC and BIC. However, due to the small sample condition of rotary lip seal, it is difficult to select the best fitting degradation process among the candidate degradation processes. Hence, IG process and Gamma process are combined by BMA method. Based on Bayesian inference, the posterior model probabilities can be inferred based on Eq. (14) and used to fuse the candidate models:

$$P\left(M_i \middle| \mathbf{Y}, \mathbf{T}\right) = \frac{P\left(M_1\right) L\left(M_1 \middle| \mathbf{Y}, \mathbf{T}\right)}{\sum_{i=1}^{2} P\left(M_i\right) L\left(M_i \middle| \mathbf{Y}, \mathbf{T}\right)} \tag{14}$$

where $L(M_i|\mathbf{Y},\mathbf{T})$ is likelihood of model M_i, $P(M_i|\mathbf{Y},\mathbf{T})$ is posterior probability of that the model M_i is the true model for degradation dataset, $\mathbf{Y}$ and $\mathbf{T}$, $P(M_i)$ is the prior for model probability of M_i. Furthermore, in the following parts of this paper, M_1 means M_{IG}, M_2 means M_{Ga}, and:

$$P\left(M_1\right) = \begin{cases} p_1 & M_{\mathrm{IG}} \text{ is true} \\ 0 & M_{\mathrm{Ga}} \text{ is true} \end{cases} \tag{15}$$

$$P\left(M_2\right) = \begin{cases} p_2 & M_{\mathrm{Ga}} \text{ is true} \\ 0 & M_{\mathrm{IG}} \text{ is true} \end{cases} \tag{16}$$

3. Failure mechanism analysis of rotary lip seal

As discussed above, seal wear is the most important failure factor for the seal during using in system and selected to be the degradation indictor in the presented research. Hence, the theoretical wear process is seemed as the theoretical degradation path of rotary lip seal. In this section, the wear simulation method for the seal, proposed in our previously published paper [11], is used and combined with moving-node updating strategy [4] to calculate the theoretical wear path. The maximum wear depth is selected as the degradation indicator of the seal, because it is much bigger than the minimum wear depth and the most significant wear is generally given with the most attention [11]. Furthermore, without the shaft-seal interference, the reverse pumping action will be destroyed and the seal will fail. Hence, the failure threshold of wear depth, D_w, is set as the initial shaft-seal interference.

3.1. Wear analysis of rotary lip seal

Based on our previously published wear simulation method and moving-node updating strategy, the main wear degradation simulation methodology is shown in Fig. 1, where h_w is wear depth, k_L is wear module, p_n is normal pressure, L_x is circumferential length of simulation space of microscale mixed lubrication model, L_y is axial length of that, U is translational speed of shaft and ω is rotational speed of shaft. Four steps need to perform to predict the theoretical wear path. First, evaluate lubrication characteristics and calculate wear module based on the microscale mixed thermo-elasto-hydrodynamic (TEHD) lubrication model. Second, simulate the normal pressure based on macroscale FE model of the seal. Third, calculate wear depth based on the above wear module and normal pressure. Finally, update the sealing lip profile and FE model. Repeat the above steps until the degradation time limitation is reached.

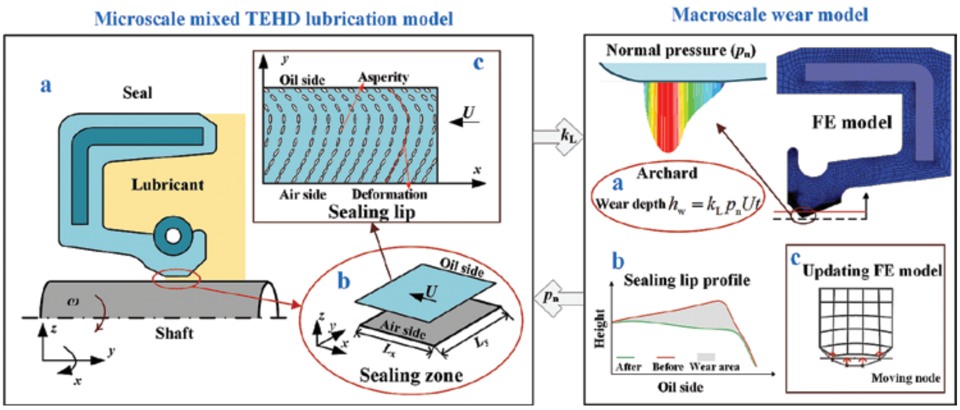


Fig. 1. Schematic diagram of wear degradation model

3.2. Theoretical degradation path and the corresponding linearization approach

Based on the above wear model and numerical simulation, the theoretical wear depth curve is predicted. By fitting the simulation results, the fitted theoretical wear path is obtained, which can be given by Eq. (17). Furthermore, the corresponding fitting results are shown in Fig. 2. It can be seen that, the theoretical wear path is not one of the common used degradation paths. Namely, even though under enough sample conditions, it is still difficult to selected the best fitting degradation path among the candidate paths based on AIC or BIC, because the theoretical wear path may not be included in the candidate paths. Hence, it is necessary to predict the wear path of rotary lip seal based on degradation mechanism analysis:

$$h_w = 54.98\ln(t+170.1) - 282.2 \tag{17}$$

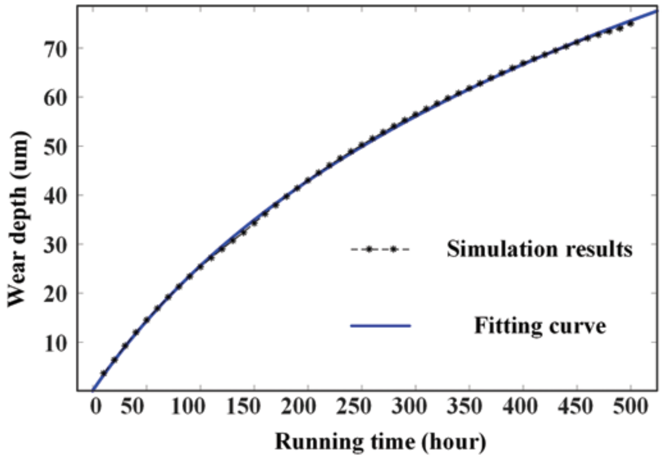


Fig. 2. Fitting results

Based on the above theoretical wear path, the linearization approach is given as follows. To transform the wear depth degradation path to a linear degradation path, the degradation of rotary lip seal is defined as:

$$y = e^{\frac{h_w+282.2}{54.98}} - 170.1 \tag{18}$$

By transforming the degradation observations, the wear depth, based on Eq. (18), the wear depth path is transformed to a linear degradation path, as:

$$\Lambda(t) = t \tag{19}$$

Furthermore, the failure threshold, D, is also transformed from the failure threshold of wear depth, D_w, based on Eq. (18). Namely:

$$D = e^{\frac{D_w+282.2}{54.98}} - 170.1 \tag{20}$$

4. Parameters estimation

Fully Bayesian inference is applied to estimate the model parameters and model probabilities. Considering small sample conditions of rotary lip seals and lack of prior related information, the priors are obtained based on MLE results for single candidate model.

4.1. Maximum likelihood estimation for IG process

Given the gradation dataset, **Y** and **T**, the corresponding likelihood function of IG process is given by:

$$L\left(\mathbf{Y}\middle|\theta_{IG}, M_{IG}, \mathbf{T}\right) = \prod_{j=1}^{n}\sqrt{\frac{\lambda_{IG} t_j^{\,2}}{2\pi y_j^3}}\exp\left[-\frac{\lambda_{IG}\left(y_j-\mu_{IG}t_j\right)^2}{2\mu_{IG}^2 y_j}\right] \tag{21}$$

where M_{IG} means IG process and $\boldsymbol{\theta}_{IG}=(\mu_{IG},\lambda_{IG})$ is corresponding parameter vector.

The corresponding log-likelihood function is given by:

$$l\left(\boldsymbol{\theta}_{IG}\middle|\mathbf{Y},M_{IG},\mathbf{T}\right)=\sum_{j=1}^{n}\left\{\frac{1}{2}\ln\left(\lambda_{IG}\right)+\ln\left(t_j\right)-\frac{1}{2}\ln\left(2\pi\right)-\frac{3}{2}\ln\left(y_j\right)-\frac{\lambda_{IG}\left(y_j-t_j\mu_{IG}\right)^2}{2\mu_{IG}^2 y_j}\right\} \tag{22}$$

The corresponding derivatives are given by:

$$\begin{cases}\dfrac{\partial l\left(\boldsymbol{\theta}_{IG}\middle|\mathbf{Y},\mathbf{T},M_{IG}\right)}{\partial\lambda_{IG}}=\dfrac{A_{IG}}{\lambda_{IG}}-\dfrac{B_{IG}}{\mu_{IG}^2}+\dfrac{C_{IG}}{\mu_{IG}}-D_{IG}\\[2ex] \dfrac{\partial l\left(\boldsymbol{\theta}_{IG}\middle|\mathbf{Y},\mathbf{T},M_{IG}\right)}{\partial\mu_{IG}}=\dfrac{B_{IG}\lambda_{IG}}{\mu_{IG}^3}-\dfrac{C_{IG}\lambda_{IG}}{\mu_{IG}^2}\end{cases} \tag{23}$$

where:

$$\begin{cases}A_{IG}=\dfrac{n}{2}\\ B_{IG}=\sum_{j=1}^{n}y_j\\ C_{IG}=\sum_{j=1}^{n}t_j\\ D_{IG}=\dfrac{1}{2}\sum_{j=1}^{n}\dfrac{t_j^2}{y_j}\end{cases} \tag{24}$$

Hence, the maximum likelihood estimation (MLE) of $\boldsymbol{\theta}_{IG}=(\mu_{IG},\lambda_{IG})$ is given by:

$$\begin{cases}\hat{\mu}_{IG}=\dfrac{B_{IG}}{C_{IG}}\\[2ex] \hat{\lambda}_{IG}=\dfrac{A_{IG}}{D_{IG}}\end{cases} \tag{25}$$

4.2. Maximum likelihood estimation for Gamma process

The likelihood function of Gamma process is given by:

$$L\left(\mathbf{Y}\middle|\boldsymbol{\theta}_{Ga},M_{Ga},\mathbf{T}\right)=\prod_{j=1}^{n}\frac{y_j^{\mu_{Ga}t_j-1}}{\Gamma\left(\mu_{Ga}t_j\right)\lambda_{Ga}^{\mu_{Ga}t_j}}\exp\left(-\frac{y_j}{\lambda_{Ga}}\right) \tag{26}$$

where θGa=(μGa,λGa) is parameter vector of Gamma process and MGa means Gamma process.

The corresponding log-likelihood function is given by:

$$l\left(\boldsymbol{\theta}_{Ga}\middle|\mathbf{Y},M_{Ga},\mathbf{T}\right)=\sum_{j=1}^{n}\left\{\left(t_j\mu_{Ga}-1\right)\ln\left(y_j\right)-\ln\left(\Gamma\left(t_j\mu_{Ga}\right)\right)-\mu_{Ga}t_j\ln\left(\lambda_{Ga}\right)-\frac{y_j}{\lambda_{Ga}}\right\} \tag{27}$$

The corresponding derivatives are given by:

$$\begin{cases}\dfrac{\partial l\left(\boldsymbol{\theta}_{Ga}\middle|\mathbf{Y},\mathbf{T},M_{Ga}\right)}{\partial\mu_{Ga}}=A_{Ga}-\sum_{j=1}^{n}t_j\psi\left(t_j\mu_{Ga}\right)-B_{Ga}\ln\left(\lambda_{Ga}\right)\\[2ex] \dfrac{\partial l\left(\boldsymbol{\theta}_{Ga}\middle|\mathbf{Y},\mathbf{T},M_{Ga}\right)}{\partial\lambda_{Ga}}=-\dfrac{B_{Ga}\mu_{Ga}}{\lambda_{Ga}}+\dfrac{2C_{Ga}}{\lambda_{Ga}^2}\end{cases} \tag{28}$$

where $\psi\left(x\right)=\dfrac{\Gamma'\left(x\right)}{\Gamma\left(x\right)}$, $\Gamma'\left(x\right)$ means derivative function of $\Gamma\left(x\right)$ and:

$$\begin{cases}A_{Ga}=\sum_{j=1}^{n}t_j\ln\left(y_j\right)\\ B_{Ga}=\sum_{j=1}^{n}t_j\\ C_{Ga}=\sum_{j=1}^{n}y_j\end{cases} \tag{29}$$

Hence, the MLE of $\boldsymbol{\theta}_{Ga}=(\mu_{Ga},\lambda_{Ga})$ can be calculated based on:

$$\begin{cases}0=A_{Ga}-\sum_{j=1}^{n}t_j\psi\left(t_j\mu_{Ga}\right)-B_{Ga}\ln\left(\lambda_{Ga}\right)\\[2ex] 0=-\dfrac{B_{Ga}\mu_{Ga}}{\lambda_{Ga}}+\dfrac{2C_{Ga}}{\lambda_{Ga}^2}\end{cases} \tag{30}$$

4.3. Priors

Based on the above MLE results, the priors of model parameters are set by Eq. (31) and Eq. (32), respectively. For the illustrative example in the presented research, the variances are set to be one-tenth of MLE results. It should be noted that the above MLE neglects the relationship between the candidate models. Hence, the above MLE results are not the final evaluation results and can be used to set the priors of model parameters.

$$\begin{cases}\pi\left(\mu_{IG}\right)\sim\mathrm{TN}\left(\hat{\mu}_{IG},0.01\hat{\mu}_{IG}^{2},0,+\infty\right),\\[2ex] \pi\left(\lambda_{IG}\right)\sim\mathrm{TN}\left(\hat{\lambda}_{IG},0.01\hat{\lambda}_{IG}^{2},0,+\infty\right)\end{cases} \tag{31}$$

$$\begin{cases}\pi\left(\mu_{Ga}\right)\sim\mathrm{TN}\left(\hat{\mu}_{Ga},0.01\hat{\mu}_{Ga}^{2},0,+\infty\right)\\[2ex] \pi\left(\lambda_{Ga}\right)\sim\mathrm{TN}\left(\hat{\lambda}_{Ga},0.01\hat{\lambda}_{Ga}^{2},0,+\infty\right)\end{cases} \tag{32}$$

where $\hat{\mu}_{IG}$ and $\hat{\lambda}_{IG}$ are the MLE results for IG process, as shown in Eq. (25). $\hat{\mu}_{Ga}$ and $\hat{\lambda}_{Ga}$ are the MLE results for Gamma process, Eq. (30). TN(·,·,0,+∞) means truncated normal (TN) distribution and the corresponding PDF is given by:

$$f_{\mathrm{TN}}\left(x\middle|\mu_x,\sigma_x,0,+\infty\right)=\begin{cases}\dfrac{\phi\left(\dfrac{x-\mu_x}{\sigma_x}\right)}{1-\Phi\left(\dfrac{-\mu_x}{\sigma_x}\right)} & 0\le x\\ 0 & \text{other}\end{cases} \tag{33}$$

The model prior probabilities generally can be determined based on expert experience and previous degradation data. However, the non-informative priors are always used under non prior information conditions. In the presented work, adopting non-informative prior, the priors of model probabilities are set as:

$$p_1=p_2=0.5 \tag{34}$$

4.4. Fully Bayesian inference

Fully Bayesian inference method is introduced to evaluate the model parameters and model probabilities, Eq. (35). The posterior distributions of model parameters and model probabilities can be obtained based on Eq. (36) using software OpenBUGS by Markov Chain Monte Carlo (MCMC) method. Furthermore, mean values of posterior distributions are used to evaluate the model parameters and model probabilities.

$$p\left(\boldsymbol{\theta},M\middle|\mathbf{Y},\mathbf{T}\right)=\frac{\sum_{k=1}^{2}L\left(\mathbf{Y}\middle|\boldsymbol{\theta}_k,M_k,\mathbf{T}\right)\pi\left(\boldsymbol{\theta}_k\right)\pi\left(M_k\right)}{\sum_{k=1}^{2}\int_{\boldsymbol{\theta}_k}L\left(\mathbf{Y}\middle|\boldsymbol{\theta}_k,M_k,\mathbf{T}\right)\pi\left(\boldsymbol{\theta}_k\right)\pi\left(M_k\right)d\boldsymbol{\theta}_k} \tag{35}$$

$$p\left(\boldsymbol{\theta},M\middle|\mathbf{Y},\mathbf{T}\right)\propto L\left(\mathbf{Y}\middle|\mu_{\mathrm{IG}},\lambda_{\mathrm{IG}},\mu_{\mathrm{Ga}},\lambda_{\mathrm{Ga}},p_1,p_2\right)\pi\left(\mu_{\mathrm{IG}},\lambda_{\mathrm{IG}},\mu_{\mathrm{Ga}},\lambda_{\mathrm{Ga}},p_1,p_2\right) \tag{36}$$

where $\boldsymbol{\theta}_1=\boldsymbol{\theta}_{\mathrm{IG}}$, $\boldsymbol{\theta}_2=\boldsymbol{\theta}_{\mathrm{Ga}}$, $\boldsymbol{\theta}=(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$. $\pi(\boldsymbol{\theta}_k)$ is the priors of model parameters of k-th candidate model and set based on the MLE results, as shown in Section 4.3. $\pi(M_k)$ is priori probability of k-th candidate model and non-informative prior is used, as shown in Eq. (34).

4.5. Estimation procedure

The parameters estimation procedure is shown in Fig. 3. Three steps need to perform to evaluate the model parameters and model probabilities. First, based on the degradation dataset, calculate MLE of each candidate model. Second, the priors are set based on the MLE results and non-informative prior. Finally, based on above priors and likelihood function, the posterior distributions of model probabilities and model parameters are obtained by fully Bayesian inference method. Furthermore, the model probabilities and model parameters are evaluated based on the mean values of the above posterior distributions. The algorithm for the case study in the presented research is as shown in Algorithm 1. n_{S} is beginning number of the transformed degradation observation and n_{E} is ending number transformed degradation observation. In the case study, nineteen degradation observation are measured. Furthermore, to demonstrate the effectiveness of the proposed method, the last four transformed degradation observations are retained as cross-validation observations. Hence, the beginning number n_{S}=15 and the ending number n_{E}=18.

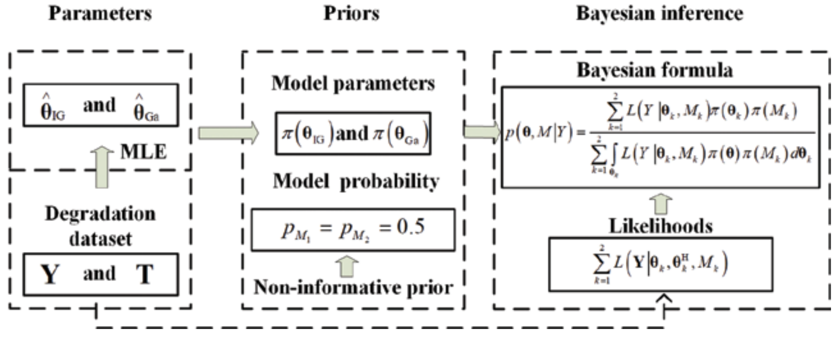


Fig. 3. Parameters estimation procedure

Algorithm 1. Simulation for case study in the presented research

Predict the theoretical wear path based on the wear analysis method, see Section 3.1.
Fit the simulation results of seal wear depth, Eq. (17)
Convert the experimental observations of wear depth based on Eq. (18).
Convert the failure threshold of wear depth to the threshold of the defined degradation based on Eq. (20).
For i=n_{S}: n_{E}
Calculate the MLE of the candidate models based on the first i transformed degradation observations, see Section 4.1 and Section 4.2.
Set the priors of the model parameters and model probabilities based on the above MLE results, see Section 4.3 and Eq. (31)-Eq (34).
Infer the model parameters and model probabilities based on the above priors and Bayesian inference, see Section 4.4.
Estimate the model parameters and model probabilities based on the mean values of the above inferring results.
Predict the degradation at the (i+1)-th predicting time.
End for
Calculate the reliability curve based on the above estimated results.
Calculate the MTTF curve based on the above estimated results.

5. Case study

In order to validate the proposed method, a typical rotary lip seal has been tested. The tested seal is as same as that presented in Ref. [12].

5.1. Experimental approach

The schematic diagram and photograph of the test bench are shown in Fig. 4. The motor is used to drive the shaft. The speed sensor is installed between the motor and shaft to measure the shaft speed. In the presented case, the shaft speed is 5000 rpm. There are two rotary lip seals are installed in the test bench. The bottom one is the tested sample. The top one is only used to seal the tank. A thermocouple is put into the sealing tank to measure oil temperature, which is controlled by adjusting the flow rates of oil in and oil out. In the presented case, the oil temperature is controlled around 25°C. Furthermore, the sealed pressure is also controlled by adjusting the flow rates of oil in and oil out. The sealed pressure is 0.2 MPa in the presented case. The sealing lip profile is measured by a profilometer. The seal wear depth is recorded every 20 hours for the first 15 observations and 50 hours for last 4 observations.

4.2. Discussions

The degradation plots of the seal wear depth by experiment are shown in Fig. 5. It can be seen that it is

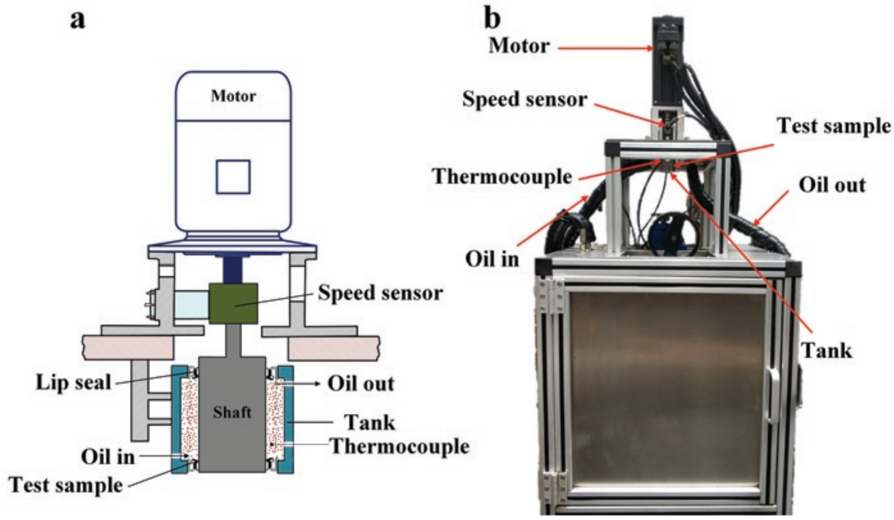


Fig. 4. (a) Schematic diagram of the test bench and (b) Photograph of the test bench

difficult to predict the degradation path of the seal wear depth. Hence, it is necessary to predict the degradation path based on the failure mechanism analysis. In the presented research, the measured seal wear depth is transformed to the defined degradation, Eq.. The transformed degradation observations are shown in Fig. 6. It can be seen that the transformed degradation observations obviously follows a linear degradation process. Namely, the theoretical wear path and the corresponding linearization method are suitable and effective. In the presented case, to validate the proposed degradation modeling method, the observations are transformed based on the above method, Eq., and are used to evaluate the model parameters and probabilities. Furthermore, to demonstrate the effectiveness of the proposed method, the last four transformed degradation observations are retained as cross-validation observations.

The prediction procedure is shown in Algorithm 1. For example, to predict the degradation at 350 hours, based on the first fifteen transformed degradation observations, for each candidate model, the parameters are evaluated by MLE. The MLE results are shown in Table 1. It should be noted that the above MLE calculation neglects the relationship between the candidate models. Hence, the above MLE results are not the final estimation results and just used to set the priors of model parameters.

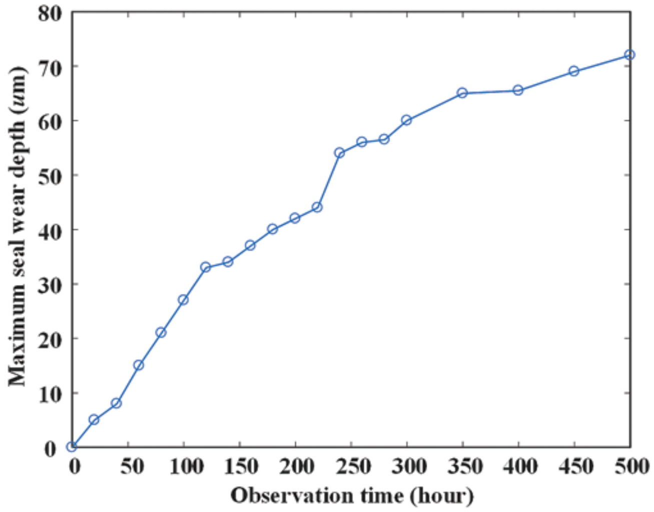


Fig. 5. Observed degradation data

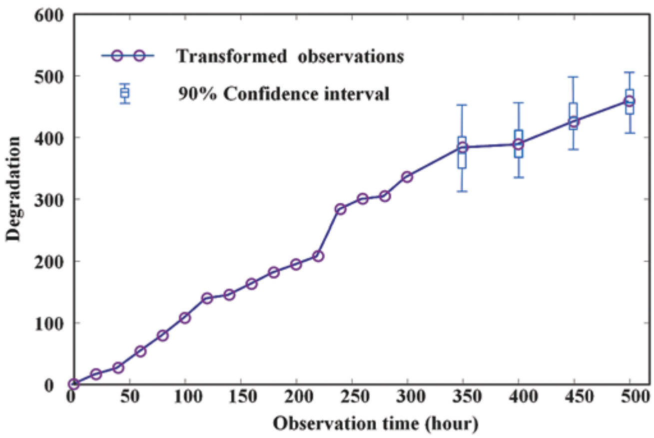


Fig. 6. Transformed degradation observation

Table 1. MLE results based on the first fifteen degradation observations

Parameter	μ_{IG}	λ_{IG}	μ_{Ga}	λ_{Ga}
MLE results	1.0563	0.0065	0.2113	2.9212

Based on the above MLE results, according to Eq. (31) and Eq. (32), the priors for the model parameters are set as Eq. (37) and Eq. (38), respectively:

$$\begin{cases}\pi(\mu_{IG})\sim\mathrm{TN}\left(1.0563, 0.10563^2, 0, +\infty\right), \\ \pi(\lambda_{IG})\sim\mathrm{TN}\left(0.0065, 0.00065^2, 0, +\infty\right)\end{cases} \tag{37}$$

$$\begin{cases}\pi(\mu_{Ga})\sim\mathrm{TN}\left(0.2113, 0.02113^2, 0, +\infty\right) \\ \pi(\lambda_{Ga})\sim\mathrm{TN}\left(2.9212, 0.29212^2, 0, +\infty\right)\end{cases} \tag{38}$$

To predict the degradation at 350 hours. Based on the above priors of model probabilities and model parameters, the model parameters and model probabilities are estimated by fully Bayesian inference method, Eq. (36), with MCMC method using OpenBUGS [13]. The posterior model probabilities are shown in Fig. 7. It can be seen that both the Gamma process and the IG process are possibly the true degradation model for the seal. Hence, it is necessary to consider the degradation process uncertainty and combine the candidate processes. Furthermore, the MCMC simulation results for model parameters are shown in Table 2. The mean values are used to evaluate the model parameters. Based on the above evaluated model probabilities and model parameters, the degradation at 350 hours can be predicted.

Furthermore, to validate the predicting accuracy of the proposed method, the last four degradation observations are retained as cross-validation observations. The prediction procedure is as same as the above predicting procedure for predicting the degradation at 350 hours. Namely, the all observations measured before the predicting time are used to estimate the model probabilities and model parameters for predicting the degradation at the predicting time. The predicting results are shown in Fig. 8. Based on the predicting results and the last five transformed degradation observations, it can be seen that

Table 2. Model parameters' MCMC simulation results based on the first fifteen degradation observations

Parameter	Mean	Standard deviation	Upper band (2.5%)	Lower band (97.5%)
μ_{IG}	1.058	0.3233	0.4266	1.692
λ_{IG}	0.02298	0.01666	0.0009685	0.06204
μ_{Ga}	0.3745	0.0584	0.2771	0.5051
λ_{Ga}	2.889	0.4354	2.094	3.798

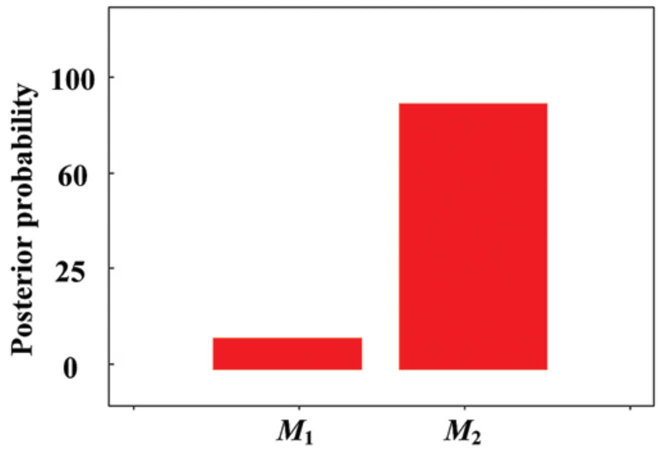


Fig. 7. Posterior probabilities of candidate models

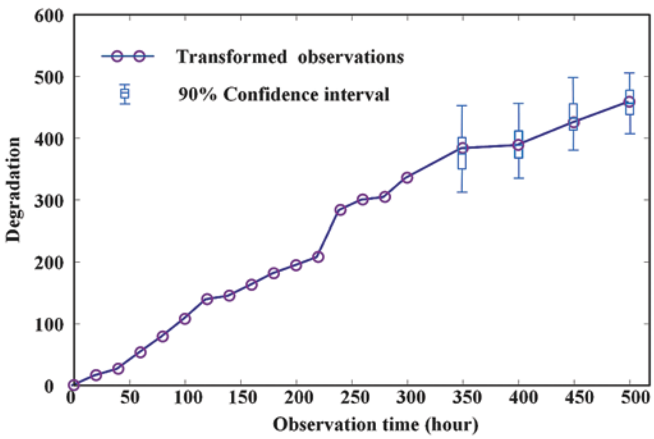


Fig. 8. Degradation prediction

the proposed method can predict the degradation of the seal precisely. Hence, the effectiveness of the proposed method has been verified.

Based on the proposed method, the reliability curve and useful lifetime distribution of the seal can also be given, Fig. 9 and Fig. 10, respectively. According to the reliability curve, the rotary lip seal can guarantee high reliability, better than 0.9, for the first 700 hours. Then the reliability of the seal quickly decreases from 700 hours to 900 hours. After 900 hours, the rotary lip seal is too unreliable to be used. According the useful lifetime distribution, it has high probability that the useful lifetime is around 800 hours. Furthermore, according to the reliability curve and useful lifetime distribution, the high reliable useful lifetime is about 750 hours for the tested rotary lip seal.

Based on the proposed method, the MTTF can also be evaluated. To investigate the sensitivity of MTTF to the failure threshold of the seal wear depth, D_w, the MTTFs under different failure thresholds are calculated and shown in Fig. 11. In the presented case, the initial seal-

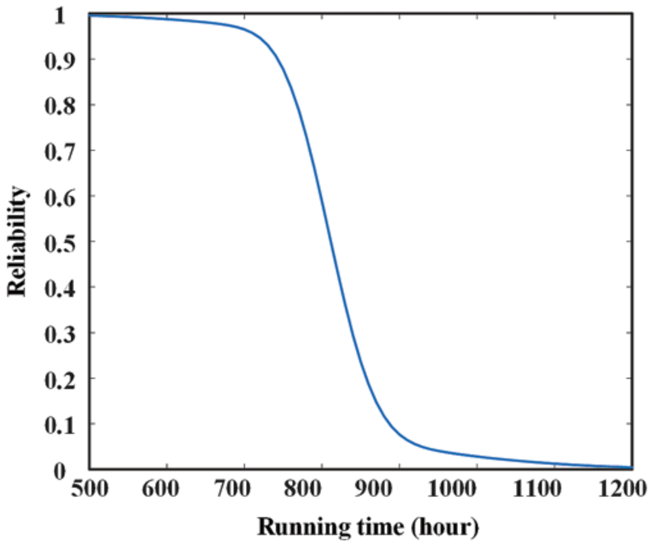


Fig. 9. Reliability curve depth

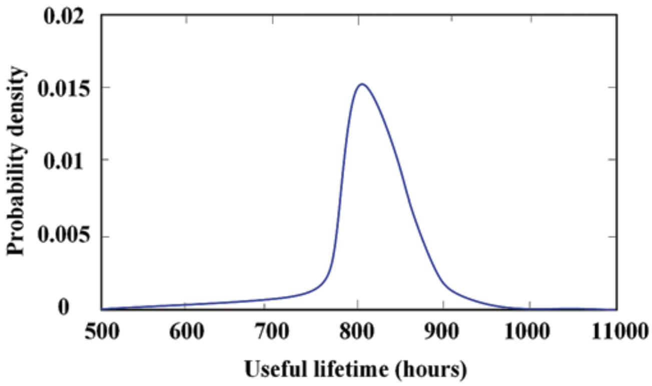


Fig. 10. Useful lifetime distribution

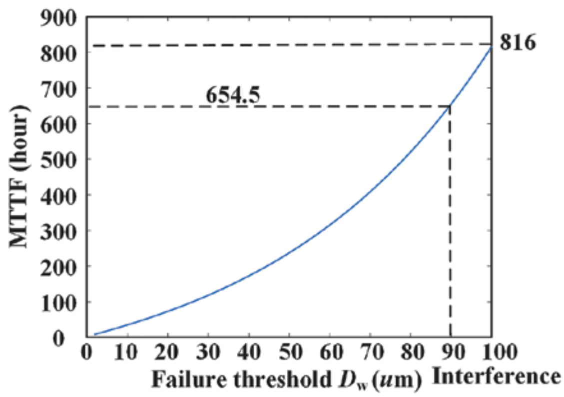


Fig. 11. Relationship between MTTF and failure threshold of maximum wear depth

shaft interference is 100um. The reverse pumping action is based on the seal-shaft interference, so the failure threshold of seal wear depth should be smaller than the initial seal-shaft interference. The MTTF is about 654.5 hours when the failure threshold is set to be 90um, while the MTTF increases to 816 hours when the failure threshold is set to be 100um. Namely, the MTTF is very sensitive to the failure threshold, especially when it approaching to the initial interference. Hence, an effective way to improve useful lifetime of the seal is to increase the failure threshold of seal wear depth, for example improving reverse pumping effects by texturing shaft surface.

6. Conclusions

In this paper, a degradation modeling method for rotary lip seal is proposed based on failure mechanism analysis and stochastic process. Stochastic process is introduced to consider randomness of rotary lip seal degradation. Considering monotonous degradation of rotary lip seal, Gamma process and IG process are selected as the candidate models and combined by BMA method. The theoretical wear path is predicted by the failure mechanism analysis and numerical simulation. A linearization method is proposed based on the theoretical wear path. In order to transform the degradation process of rotary lip seal to a linear process, the measured degradation data is transformed based on the linearization method. The model parameters and probabilities are evaluated by fully Bayesian inference method. The effectiveness of the proposed method has been verified by comparing the predicting degradation and experimental observations.

Based on the proposed method, the MTTF, reliability curve and useful lifetime distribution of the tested rotary lip seal can also be evaluated. The sensitivity of MTTF to the failure threshold of seal wear depth is analyzed. It can be seen that the MTTF is very sensitive to the failure threshold, especially when it is approaching to the initial interference. Hence, an effective way to improve useful lifetime of the seal is to increase the failure threshold of the seal wear depth, for example improving reverse pumping effect by texturing shaft surface.

The presented work provides a foundation to evaluate the useful lifetime and reliability of rotary lip seals. Future works will focus on how to improve the reverse pumping action and seal wear by texturing the shaft surface.

Acknowledgement

This study was co-supported by the National Natural Science Foundation of China (51620105010, 51575019), Natural Science Foundation of Beijing Municipality (17L10011), and Program 111.

References

1. Adrian ER, David M, Jennifer A. Bayesian Model Averaging for Linear Regression Models. Journal of the American Statistical Association 1997; 92: 179-91, https://doi.org/10.1080/01621459.1997.10473615.
2. Cao W, Jia X, Liu Y, Hu QW, Zhao JM. Selective maintenance optimisation considering random common cause failures and imperfect maintenance. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2019; 233(3): 427-443, https://doi.org/10.1177/1748006X18799907.
3. Chen Q, Wang S, Liu D, Zhang C. Numerical analysis of wear degradation model of rotary lip seal with multifield coupling. Proceedings of IEEE 8th International Conference on Fluid Power and Mechatronics; 2019 Apr. 10-13; Wuhan, China. New York: IEEE; 2019; p. 1-6.
4. Frölich D, Magyar B, Sauer B. A comprehensive model of wear, friction and contact temperature in radial shaft seals. Wear 2014; 311(1-2): 71-80, https://doi.org/10.1016/j.wear.2013.12.030.
5. Gadari ME, Fatu A, Hajjam M. Shaft roughness effect on elasto-hydrodynamic lubrication of rotary lip seals: experimentation and numerical simulation. Tribology International 2015; 88(2015):218-27, https://doi.org/10.1016/j.triboint.2015.03.013.
6. Guo F, Jia X, Lv M, Wang L, Salant RF. The effect of aging in oil on the performance of a radial lip seal. Tribology International 2014; 78: 187-94, https://doi.org/10.1016/j.triboint.2014.05.017.
7. Guo F, Jia X, Longke W. The effect of wear on the performance of a rotary lip seal. Journal of Tribology 2014; 136(4): 041703, https://doi.org/10.1016/j.triboint.2014.05.017.
8. Guo F, Jia X, Huang L, Salant RF. The effect of aging during storage on the performance of a radial lip seal. Polymer Degradation and Stability 2013; 98(11): 2193-200, https://doi.org/10.1016/j.triboint.2014.05.017.
9. Li G, Zhang Q, Huang E, Lei Z, Wu H, Xu G. Leakage performance of floating ring seal in cold/hot state for aero-engine. Chinese Journal of Aeronautics 2019, In press, https://doi.org/10.1016/j.cja.2019.03.004.
10. Liu D, Wang S, Zhang C, Tomovic M. Bayesian model averaging based storage lifetime assessment method for rubber sealing rings. Advances in Mechanical Engineering 2019; 11(5): 1687814019853351, https://doi.org/10.1177/1687814019853351.
11. Liu D, Wang S, Zhang C. A multiscale wear simulation method for rotary lip seal under mixed lubricating conditions. Tribology International 2018; 121: 190-203, https://doi.org/10.1016/j.triboint.2018.01.007.
12. Liu D, Wang S, Tomovic M, Zhang C. Numerical study of the effects of textured shaft on the wear of rotary lip seals. Tribology International 2019; 138: 215-38, https://doi.org/10.1016/j.triboint.2019.05.037.
13. Liu D, Wang S, Zhang C, Tomovic M. Bayesian model averaging based reliability analysis method for monotonic degradation dataset based on inverse Gaussian process and Gamma process. Reliability Engineering & System Safety 2018; 180: 25-38, https://doi.org/10.1016/j.ress.2018.06.019.
14. Liu W, He G. Storage life of silicone rubber sealing ring used in solid rocket motor. Chinese Journal of Aeronautics 2014; 27(6): 1469-76, https://doi.org/10.1016/j.cja.2014.10.013.
15. Müller HK. Concepts of sealing mechanism of rubber lip type rotary shaft seals. In: Proceedings of the BHRA 11th international conference on fluid sealing. Cannes. 1987: 698-709.
16. Park C, Padgett WJ. Stochastic degradation models with several accelerating variables. IEEE Trans Rel 2006; 55 :379-90, https://doi.org/10.1109/TR.2006.874937.
17. Park C, Padgett WJ. Accelerated degradation models for failure based on geometric Brownian motion and Gamma processes. Lifetime Data Anal 2005; 11: 511-27, https://ieeexplore.ieee.org/document/1638421.
18. Park I, Grandhi RV. A Bayesian statistical method for quantifying model form uncertainty and two model combination methods. Rel Eng Syst Safety 2014; 129: 46-56, https://doi.org/10.1016/j.ress.2014.04.023.
19. Peter JH, Malik A S, Langer K. An efficient reliability-based simulation method for optimum laser peening treatment. Journal of Manufacturing Science and Engineering 2016; 138(11): 111001, https://doi.org/10.1115/1.4033604.
20. Peng C, Guo S, Ouyang X, Zhou Q, Yang H. An eccentric 3-D fluid-structure interaction model for investigating the effects of rod parallel offset on reciprocating-seal performance. Tribology International 2018; 128: 279-90, https://doi.org/10.1016/j.triboint.2018.07.028.

21. Peng C, Ouyang X, Zhu Y, Guo S, Zhou Q, Yang H. Investigation into the influence of stretching on reciprocating rod seals based on a novel 3-D model vs axisymmetric model. Tribology International 2018; 117: 1-14, https://doi.org/10.1016/j.triboint.2017.06.020.
22. Peng CY. Inverse Gaussian processes with random effects and explanatory variables for degradation data. Technometrics 2015; 57: 100-11.
23. Peng W, Liu Y, Zhang X, Huang HZ. Sequential preventive maintenance policies with consideration of random adjustment-reduction features. Eksploatacja i Niezawodnosc -Maintenance and Reliability 2015, 17(2): 306-313, https://doi.org/10.1080/00401706.2013.879077.
24. Peng W, Li YF, Yang YJ, Mi J, Huang HZ. Bayesian Degradation Analysis With Inverse Gaussian Process Models Under Time-Varying Degradation Rates. IEEE Transactions on Reliability 2017; 66: 84-96, https://ieeexplore.ieee.org/document/7803533/.
25. Peng W, Li YF, Yang YJ, Huang HZ, Zuo MJ. Inverse Gaussian process models for degradation analysis: a Bayesian perspective. Reliab Eng Syst Saf 2014; 130: 175-89, https://doi.org/10.1016/j.ress.2014.06.005.
26. Qin H, Zhang S, Zhou W. Inverse Gaussian process-based corrosion growth modeling and its application in the reliability analysis for energy pipelines. Front. Struct. Civil Eng 2013; 7: 276-87.
27. Raftery AE, Gneiting T, Balabdaoui F, Polakowski M. Using Bayesian model averaging to calibrate forecast ensembles. Month Weather Rev 2005; 133: 1155-74.
28. Rodríguezpicón LA, Rodríguezpicón AP, Méndezgonzález LC, Rodríguez-Borbón MI, Alvarado-Iniesta A. Degradation modeling based on gamma process models with random effects. Commun Stat 2017; 1: 1-15.
29. Salant RF, Flaherty AL. Elastohydrodynamic analysis of reverse pumping in rotary lip seals with microasperities. Journal of Tribology 1995; 117(1): 53-9.
30. Shen D, Salant RF. Elastohydrodynamic analysis of the effect of shaft surface finish on rotary lip seal behavior. Tribology Transactions 2003; 46(3): 391-6.
31. Shen D, Salant RF. A transient mixed lubrication model of a rotary lip seal with a rough shaft. Tribology Transactions 2006; 49(4): 621-34.
32. Tseng ST, Balakrishnan N, Tsai CC. Optimal step-stress accelerated degradation test plan for gamma degradation processes. IEEE Transactions on Reliability 2009; 58: 611-18.
33. Wang HK, Li YF, Liu Y, Yang YJ, Huang HZ. Remaining useful life estimation under degradation and shock damage. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2015; 229(3): 200-208.
34. Wang LZ, Pan R, Li XY, Jiang TM. A Bayesian reliability evaluation method with integrated accelerated degradation testing and field information. Reliability Engineering & System Safety 2013; 112: 38-47.
35. Li X, Peng GL, Wang Q. A numerical analysis method of hydraulic seals for downhole equipments. Advances in Mechanical Engineering 2013; 5: 151794.
36. Li X, Peng GL, Li Z. Prediction of seal wear with thermal-structural coupled finite element method. Finite Elements in Analysis and Design 2014; 83: 10-21.
37. Yang Y J, Huang HZ, Liu Y, Zhu SP, Peng WW. Reliability analysis of electrohydraulic servo valve suffering common cause failures. Eksploatacja i Niezawodnosc -Maintenance and Reliability 2014, 16(3): 354-359.
38. Zio E. Some challenges and opportunities in reliability engineering. IEEE Transactions on Reliability 2016; 65(4): 1769-82.

Di LIU
Shaoping WANG
School of Automation Science and Electrical Engineering
Beihang University
Xueyuan Road, Haidian District, Beijing, P.R. China

Mileta M. TOMOVIC
Department of Engineering Technology
Old Dominion University
Norfolk, Virginia, 23529, USA

Emails: liudi54834@buaa.edu.cn, shaopingwang@buaa.edu.cn, MTomovic@odu.edu

Article citation info:
PAWEŁCZYK M, FULARA S, SEPE M, DE LUCA A, BADORA M. Industrial gas turbine operating parameters monitoring and data-driven prediction. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 391–399, http://dx.doi.org/10.17531/ein.2020.3.2.

Maciej PAWEŁCZYK
Szymon FULARA
Marzia SEPE
Alessandro DE LUCA
Maciej BADORA

INDUSTRIAL GAS TURBINE OPERATING PARAMETERS MONITORING AND DATA-DRIVEN PREDICTION

MONITOROWANIE ORAZ BAZUJĄCA NA DANYCH PREDYKCJA PARAMETRÓW ROBOCZYCH PRZEMYSŁOWEJ TURBINY GAZOWEJ

The article reviews traditional and modern methods for prediction of gas turbine operating characteristics and its potential failures. Moreover, a comparison of Machine Learning based prediction models, including Artificial Neural Networks (ANN), is presented. The research focuses on High Pressure Compressor (HPC) recoup pressure level of 4th generation LM2500 gas generator (LM2500+G4) coupled with a 2-stage High Speed Power Turbine Module. The researched parameter is adjustable and may be used to balance net axial loads exerted on thrust bearing to ensure stable gas turbine operation, but its direct measurement is technically difficult implicating the need to indirect measurement via set of other gas turbine sensors. Input data for the research have been obtained from BHGE manufactured and monitored gas turbines and consists of real-time data extracted from industrial installations. Machine learning models trained using the data show less than 1% Mean Absolute Percentage Error (MAPE) as obtained with the use of Random Forest and Gradient Boosting Regression models. Multilayer Perceptron Artificial Neural Networks (MLP ANN) models are reviewed, and their performance checks inferior to Random Forest algorithm-based model. The importance of hyperparameter tuning and feature engineering is discussed.

Keywords: *gas turbine, machine learning, data-driven prediction, HP recoup pressure analysis.*

W artykule przedstawiono przegląd klasycznych i aktualnych metod przewidywania parametrów operacyjnych oraz potencjalnych usterek turbin gazowych. Dodatkowo zaprezentowano porównanie wybranych modeli opartych o uczenie maszynowe, w tym modeli wykorzystujące sztuczne sieci neuronowe. Przeprowadzone badania dotyczyły analiz poziomu ciśnienia ze sprężarki turbiny gazowej LM2500 czwartej generacji (LM2500+G4) połączonej z dwustopniową turbiną roboczą. Badany parametr podlega sterowaniu i może posłużyć do wyrównania sił osiowych działających na łożysko główne wału wysokiego ciśnienia w celu zapewnienia stabilnej i niezawodnej pracy turbiny gazowej. Jednocześnie jego bezpośredni pomiar jest kosztowny stąd potrzeba dokonania pośredniego pomiaru z wykorzystaniem innych czujników zamontowanych na turbinie. Dane wejściowe do analiz otrzymano dzięki uprzejmości producenta turbin, firmy BHGE. Zawierają one parametry bezpośrednio pobrane z monitorowanych turbin gazowych. Modele uczenia maszynowego otrzymane w wyniku analizy charakteryzują się średnim błędem procentowym (MAPE) na poziomie poniżej 1%. Najmniejszym błędem charakteryzują się modele otrzymane przy zastosowaniu lasów losowych (Random Forest) oraz gradientowego wzmacniania regresji (Gradient Boosted Regression). Przetestowano także zastosowanie wielowarstwowych, w pełni połączonych sztucznych sieci neuronowych, których efektywność okazała się niższa od modelu opartego o algorytm lasów losowych. W podsumowaniu podkreślono wagę dostosowywania hiperparametrów i inżynierii cech.

Słowa kluczowe: *turbina gazowa, uczenie maszynowe, predykcja bazująca na danych, analiza pomiarów ciśnienia sprężarki.*

1. Introduction

Operational reliability of complex mechanical energy generation system is a key in assuring stable and cost-effective power supply for long-term commercial, industrial and communal purposes. One of the key systems used to deliver energy is currently, and for the foreseeable future, gas turbine. Gas turbine in-service monitoring proved to be useful in potential failure diagnosis and prevention as well as in its emission monitoring [36]. Appropriate gas turbine utilization, for example in de-rated temperature mode, may have significant contribution to decrease in harmful exhaust emissions [20] as well as to extend reliable operation of the system [1]. In this article accumulated gas turbine operating service data has been studied to predict key parameters.

Prediction of gas turbine performance is one of the key factors researched. In-service performance deterioration and system unreliability are significant contributors to a gas turbine utilization planning. System failure or even single percent of power reduction can result in significant impact throughout life cycle. Overhaul is the optimum remedy to restore desired parameters, but for the planning purposes, it should also be modelled and postponed if possible. Moreover, performance deterioration (and system failure) proved to be difficult to monitor and predict using installed gas turbine due to sensor accuracy, instrumentation aging, assembly constraints and control challenges. Therefore, maintenance planning should be based on multiple parameter's analysis and results integration using specialized model [12].

There are several typical examples of monitoring systems, that have been studied in the past. Historically, standard gas turbine tests have been used to create performance maps and model typical gas turbine behavior. Instrumentation proved to be one of the most important challenges, limiting data acquisition capability [11]. Gas turbine integrity issues, cost of engine testing and limited acquisition capabilities are responsible for this restraint [35]. Moreover, tests of the heavyweight gas turbines (75 MW and more) are not carried out before gas turbine delivery due to cost and assembly reasons.

Multiple statistical methods have been developed over time to preview, model and decrease gas turbine life cycle cost. Traditional analytical methods include, among others: Failure Mode and Effects Analysis (FMEA), Fault Tree Analysis (FTA), and Markov Analysis. An example of the Failure Mode and Effects Analysis (FMEA) supplemented by statistical failure rate model has been discussed by Mingazov and Korobitsin [29]. The results of such model application may allow to assess different gas turbine systems reliability. Non-traditional reliability monitoring methods and novel approaches to reliability monitoring include vague lambda-tau methodology [23], fuzzy sets [37] and numerical assessment based on Piecewise Deterministic Markov Process and Quasi Monte Carlo methods [10]. The Petri Net model fed by Failure Tree Analysis (FTA) has also been presented and discussed by Verma and Kumar [35]. To address system unreliability, the fuzzy sets application has been studied by several authors, example of which is described by Huang et all [18]. The results obtained using these methods are promising, however they also require rough estimation of expected reliability to proactively choose appropriate modelling method. The dynamic nature of gas turbine can also be taken into consideration. The power generation system deteriorates over time and thus its reliability is time dependent. Binary Decision Diagram (BDD) has been proposed to address a multi-phase network system as discussed by Lu et al. [26]. The industrial gas turbine system operation may also be considered as phased-mission system (PMS) since their operation mode usually consists of numerous repetitive phases. Therefore, a flexible truncation limit may be applied to BDD problem explosion. While the truncation application is necessary to model increasing number of phases, the flexible truncation allows retention of the truncation error at low level as presented by Lu et al. in [27].

The statistical-based analytical models are the most frequently used methods for Engine Health Monitoring (EHM). Trend analysis is one of the standard methods used to compare current gas turbine parameters with warning levels [32]. Pattern recognition can be employed in limited time frames (or sliding windows) to detect gas turbine behavior anomaly and potential failure. Kalman filters are applied to define parameter changes that result in least square error. Artificial Neural Networks (ANN), support vector machines (SVM) and particle swarm optimization (PSO) techniques are also utilized to improve the accuracy of the predictions [6,7,16,21,33]. It has been proven that a hybrid PSO-SVM based model can result in a regression accuracy of approximately 95% [13]. It is also worth noting that described methods can achieve top results only when applicable dataset is available in vast quantities for multiple sensors (implicating Big Data solutions). In general, it can be concluded that the physical models derived from full or partial operating parameters provide the most accurate results [22, 24].

As discussed, performance optimization and hardware ageing modelling are crucial to maintain high efficiency of the propulsion system and to monitor health condition of the critical components. Usually, only selected parameters are remotely monitored. The limited data acquisition requirement allows minimization of costs associated with the installation of additional sensors in hardly accessible turbine locations. In addition, the benefit of installing advanced sensors to monitor system performance and components health status must be balanced against the risk of gas turbine downtime due to sensor malfunctions and potential secondary damages of gas turbine components caused by sensor failure itself. A scenario of gas turbine life-cycle optimization with several operational parameters used as input to advanced models estimating online gas turbine performance and components ageing has been presented by Baker Hughes team [28]. The team concluded that algorithm's prediction accuracy together with data quality and proper expertise are important to model accurately long-term gas turbine behavior and to provide correct maintenance insights to onsite operators [28].

To ensure consistently high accuracy required, data exploration of each input and corresponding data processing is required [17]. While main gas turbine control parameters are redundant and transducer outputs are processed by the control software according to robust selection criteria and technical board approved logic, some parameters are obtained directly from a single sensor and hence a predictive modelling may help to create a baseline for consistency check.

2. Presentation of the problem

This paper discusses the use of machine learning (ML) algorithms to estimate chosen parameter (in this case, the secondary flow pressures, particularly the HP recoup pressure) with respect to the main gas turbine control parameters in order to troubleshoot instrumentation anomalies or detect deviations from the expected baseline. The primary goal is to predict long-term HP recoup pressure and hence enable consistent performance prediction and potential early detection of components degradation.

The HP recoup pressure is an adjustable parameter, that enables balancing of net axial loads exerted on thrust bearing and ensures operation within desired threshold. The control method and applicable system have been described and patented by Badeer [2].

The presented research is based on data obtained from 4th generation LM2500 gas generator (LM2500+G4) coupled with a 2-stage High Speed Power Turbine Module GE Challenge competition [19]. A cross section of LM2500 gas turbine is shown on Fig. 1. The LM2500 is an industrial derivative of the General Electric (GE) CF6 aircraft engine (engine originally developed for aviation purposes and then refurbished for stationary operation) [3, 14]. This study focuses on PGT25+G4 gas turbine delivering 34 MW with thermal efficiency of 41% [5].

The presented study shows analytical comparison of the gas turbine engine pressure parameter prediction with respect to other operational (numerical) and geometrical (categorical) parameters. Categorical parameter is represented by separately provided pressure orifice size, available only for limited number of gas turbines. The orifice plate is built in the HP Recoup pressure line to allow flow rate measurement and is causing an irrecoverable pressure loss [30]. The

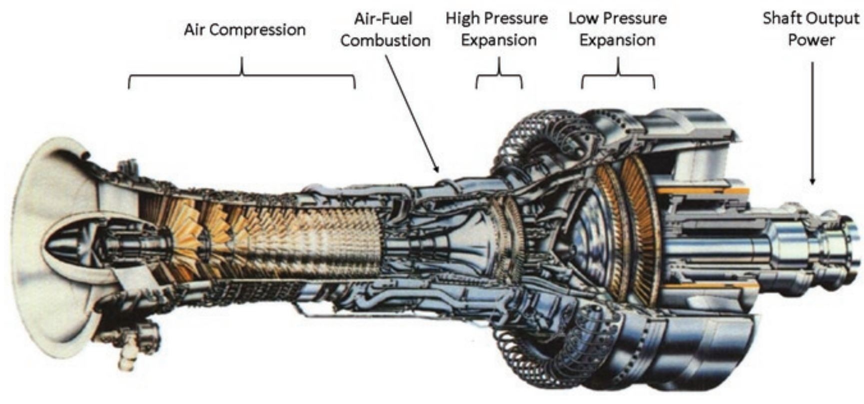


Fig. 1. Researched industrial & marine gas turbine configuration, based on [8]

potential impact of the orifice size on the researched HP Recoup pressure parameter prediction has been also studied.

3. Mathematical modelling and models implementation

Since turboshaft engine consists a close-loop system, by design, researched recoup pressure is mutually depended on other turbine engine parameters. This feature makes it suitable for multistep prediction, as the best model can be identified in such condition [31].

3.1. Mathematical models

Typical starting point for all machine learning challenges are linear models. This include linear regression (logistic regression in instance of models with categorical values included) with linear coefficients κ = (κ_1, κ_2, …, κ_n) and biases assigned to each of the input variables to minimize the output sum of squares between the training labels and predictions via linear approximation:

$$\hat{\kappa} = \min_{\kappa} X\kappa - y_2^2 \tag{1}$$

where $X\kappa = y$ is linear equation with coefficient κ and $\hat{\kappa}$ is predicted value.

While continuous numerical inputs for the model are normalized (as discussed later in this article) to prevent data range discrepancies from affecting the resultant model, additional options are available to prevent high coefficients from being assigned to selected parameters, while keeping the others small (and effectively eliminating them from the prediction process). This is implemented by imposing additional penalty parameter on the size of coefficient in the cost function calculation process. For ridge regression (also known as Tikhonov based regression or L2) the penalty is assigned to a sum of squared weights. For LASSO (Least Absolute Shrinkage and Selection Operator or L1) regression penalty value is only assigned to the sum of absolute values of the model weights (which can allow removal of redundant feature from the model), while elastic net model allows utilization of both routines at the same time. This process forces linear model weight to remain small, which typically leads to less overfitting [34]. The penalty value is tuned to obtain minimum testing set prediction error:

$$\hat{\kappa} = \min_{\kappa} \frac{1}{2m_j} \|X\kappa - y\|_2^2 + \alpha \|\kappa\|_1 \tag{2}$$

where m_j is a number of samples, α is constant and $\|\kappa\|_1$ is ℓ_1-normal to coefficient κ.

LASSO/ridge/elastic net CV models have additional cross validation routines built into them to automatically split the entire dataset into all possible train/test combinations to allow model generalization on the entire dataset, while maintaining dataset split. This happens on the expense of model performance as cross validation split of 5 will require training of 5 models, rather than one. LassoLarsIC is computationally cheaper alternative to this process.

Bayesion Ridge Regression models provides a probabilistic estimate for a regression problem. The model utilizes multiple parameters to obtain similar regularization pattern to ridge regression. The parameters are estimated by maximizing the marginal log likelihood. Bayesian Ridge Regression provides a different weight set than those obtained using Ordinary Least Squares method:

$$p(\kappa|\lambda) = \mathcal{N}\left(\kappa|0, \lambda^{-1}\mathbb{I}_p\right) \tag{3}$$

where λ is Gaussian distribution for probabilistic model of output parameter κ.

Kernel models allow for dataset dimensionality modification/reduction, which allows model to create an artificial hyperspace split between model parameters, which can result in better model performance. This method can be applied on models such as Support Vector Machine (SVM), which can be used for either classification (Support Vector Classifier or SVC) or regression (Support Vector Regression) tasks. SVM works by finding a set of optimum linear split in high dimensional space, using a subset of training points in the decision function (support vectors). While different kernels can be used to achieve optimum model performance nonlinear kernels are computationally expensive, especially for large datasets.

Random Forest Model consists of multiple decision trees (with exact decision tree numbers determined based on hyperparameter tuning), which are averaged to classify target feature properly. During construction of each tree a node split is randomly generated to best fit random sub-sample of the features. The randomness results in increase of forest classifier bias, but, due to averaging, the model variance also decreases. This compensates the adverse effect leading to higher quality prediction. The algorithm processes original dataset by proposing a randomly generated shallow trees, usually bootstrapping (randomly sampling with replacement) original dataset for to augment original training dataset. Once trained, a prediction is made for every tree in the forest and individual estimates are combined to achieved optimum testing set performance. For regression problem (no categorical feature included) the overall result is a mean of individual predictions. For classification problem (orifice size included) each tree provides a weighted confidence for each class, which are then averaged across all trees. The result is the class with the highest confidence to minimize general mean-squared generalization error L for the numerical prediction $\hat{y}_i$ and its corresponding true label y_i:

$$L = \sum_{i=1}^{N} \left(y_i - \hat{y}_i\right)^2 \tag{4}$$

The binary nature of the Random Forest Model can be partially compensated by iteratively reducing the error rate for each of the created trees, by building regression tree on negative gradient of a loss function, quite similarly to neural network training process. This method is called gradient boosted decision trees (GBDT) [25,38]. It works by processing a small number of decision trees (which in this instance are base learners) and using a learning rate to adjust tree definition to minimize aggregation of the loss function $L(y_i, F(x_i))$, by reducing individual tree error:

$$T^* = arg \min_{T} \sum_{i=1}^{N} L\left(y_i, F\left(x_i\right)\right) \tag{5}$$

Artificial Neural Networks (ANN) model training process, often called Deep Learning (for 3 hidden layers or more). Typical neural network relies on a predefined input and output vector and inner stacked layers utilizing mathematical transformations of input data as shown on Fig. 2. ANNs work by processing input vector through a set of linear functions followed by nonlinear activation (recently almost popular being Rectified Linear Unit or ReLU) in a process called feed forward propagation. This is followed by loss (for a single batch) and cost (for entire dataset) value computation based on predefined cost function. Then, for each weight, a gradient with respect to specific parameter is computed using a chain rule and cost value. Finally, all weights are updated sequentially using process called backpropagation.

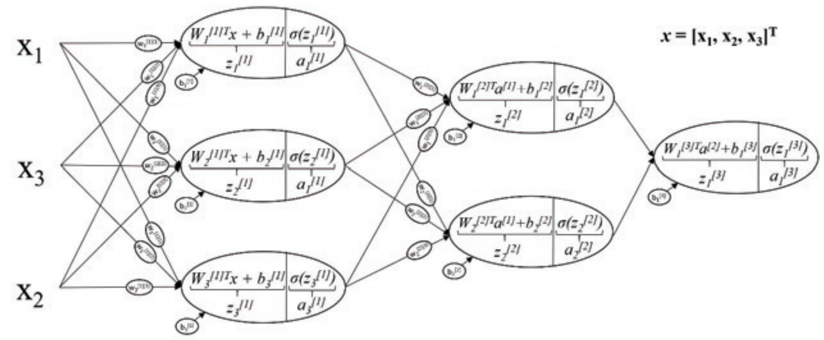


Fig. 2. Overview of Artificial Neural Network (ANN) densely connected architecture with 2 hidden layers

Considering that deep neural networks, especially convolutional neural networks for image processing, (utilizing a combination of convolutional, pooling and dense neural networks) usually consist of 10+ layers, it is useful to represent given neural network via high level representation. The example of 3 hidden layer network with 4 nodes in each layer, 3 inputs and 2 outputs has been shown on Fig. 3.

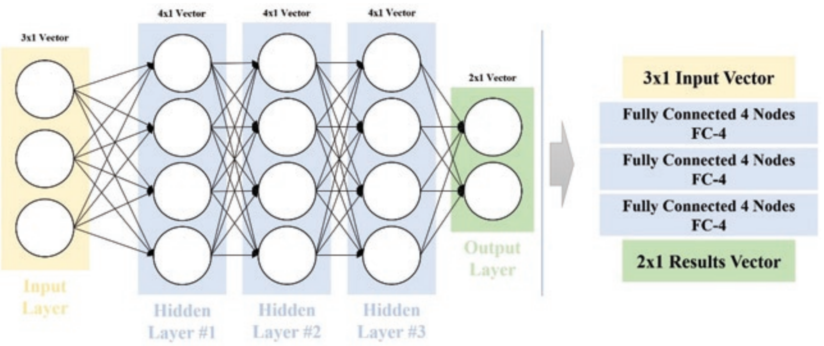


Fig. 3. An example high level representation on densely connected neural network. For the purpose of analyzed research 3 hidden layer network with up to 20 inputs, up to 100 nodes for each layer and 1 output has been used

This high level, generalized representation is often used in Deep Learning frameworks such as Keras, PyTorch. While Scikit-learn offers simple neural network implementations, those frameworks, while working on top of low level backends (Tensorflow), offer much more flexibility and significant reduction in training and inference (model implementation) time, typically due to effective processing routines for direct implementation on CPU/GPU (and recently, TPU).

3.2. Models Implementation

Models obtained in the following analysis were developed using Python 3.6, modules such as Numpy/Pandas (data processing), Scikit-learn (machine learning) and Keras/Tensorflow (Artificial Neural Networks) frameworks. Model was iteratively developed, with certain analysis thresholds established through set of tests.

In order to properly establish acceptability threshold for model's results a benchmark method has been proposed. It consists of three metrics used to compare results of applied machine learning models: Mean Absolute Percentage Error (MAPE), Root-Mean-Square Error (RMSE) and Coefficient of Determination (R2). The three parameters have been chosen as they represent a basis for qualitative and quantitative comparison of the researched methods. MAPE metric allows valuable insight into the absolute magnitude of the error, while RMSE provides additional intuition into scaled residual size. R2 score represents a statistical measurement of model fit into predicted dataset. Target feature mean has been assumed as baseline prediction, resulting in RMSE of 0.83, MAPE of 9.24% and R2 score of 0 (per R2 score definition).

The first step in predictive modelling is data exploration, which allowed to limit the number of features used as predictive maintenance system input. Data cleaning pipeline implementation allowed dataset feature size reduction from 4568 columns to 382. Any parameter with more than 3% missing data has been removed. For remaining parameters missing data has been replaced with median for all other features. Additionally, all outliers (exceeding +/− 6 Sigma deviation) were removed (entire data row removal) to prevent errors further on in the analysis.

Furthermore, categorical geometry feature – flow orifice size – has been provided and added to the dataset. Subsequent analysis has shown that geometry have significant predic-

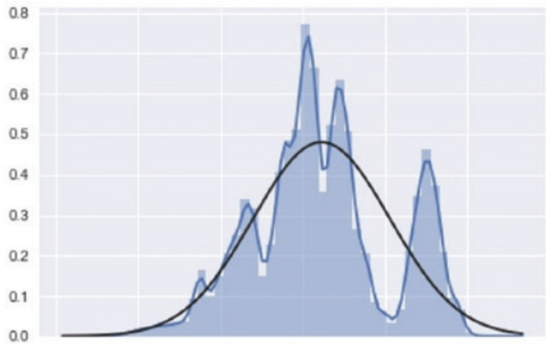


Fig. 4. Distribution of the target feature (gas turbine pressure) with fitted normal distribution. As input dataset (after initial preprocessing) consisted of multiple serial numbers, the resultant dataset is multinomial. No ensemble subgrouping has been used due to sufficient model predictive capacity

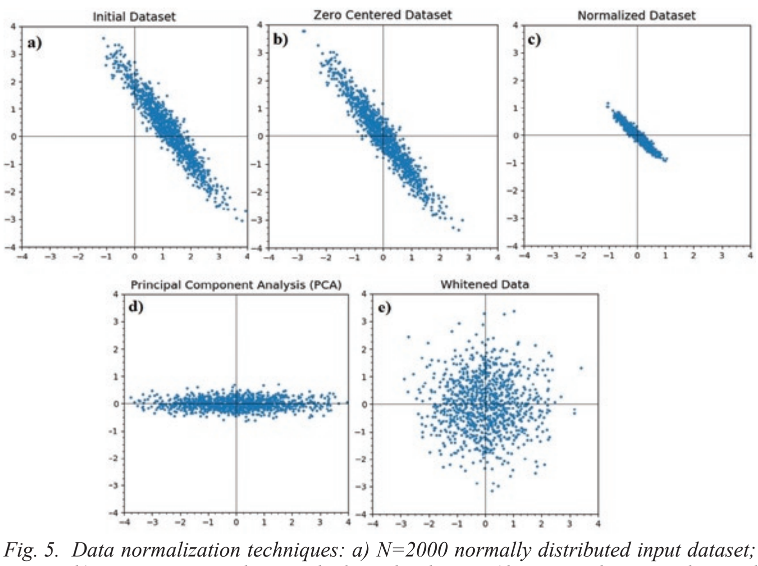


Fig. 5. Data normalization techniques: a) N=2000 normally distributed input dataset; b) centering procedure applied on the dataset (dimensional mean subtracted from each datapoint); c) Standard deviation-based normalization of centered data; d) Principal Component Analysis (PCA) applied on the initial dataset; e) Whitening transformation applied on input dataset

tive power and as such all rows, where geometrical data was not available, were dropped. Resultant target feature (pressure) distribution is shown on Fig. 4.

High predictive power of the model has been ensured by ranking remaining features based on Pearsons's correlation coefficient calculated between each of 380 analyzed parameters and target feature. Feature independence has been ensured by analyzed the relationship between pairs and picking only sensors, which would allow physics-based explanation and higher correlation value with target feature. This approach has been chosen (instead of proceeding with typical correlation tests like Pearson's or Chi-Squared) to streamline dataset reduction through project approval board. This allowed reduction of input dataset to 45 features. As such remaining 45 features were additionally tested with random forest feature estimator. Combined feature mark considered both regression correlation and random forest estimator mark. Best 20 features were used in the further analysis and combined with geometry categorical feature. Random 5% of the dataset were retained as model test set.

Input data set could be normalized in order to allow more effective computation paradigm and better convergence for neural network processing. Typical normalization steps consist of combination of zero-centering/normalization. Similarly, Principal Component Analysis (PCA) is used as an efficient way of both data normalization and dimensionality reduction. Finally, data whitening transformation ensures that mean data mean value is concentric with coordinate system origin point as well as that all the axis in the dataset have the same variance. These normalization methods have been presented on Fig. 5. Due to variable nature of turbine operating parameters input dataset has not been normalized. This approach has been tested on validation dataset and has shown more consistent results then normalizing input data.

For each of the predefined model combination data has been preprocessed to better fit specified model, then a test/train split has been created using random number seed and test set size of 0.25 (25% of the dataset was retained for testing) followed by classifier training, model storage and result view for postprocessing. Example of Linear Regression model implementation is shown on Figure 6. It shows Linear Regression model implementation.

```
modelNames = [["Linear Regression With Orifice", "WO", "sklearnRegression"],
              ["Linear Regression with No Orifice", "NO", "sklearnRegression"]]

for modelNameIndex, (modelName, orificeStatus, saveName) in enumerate(modelNames):

    print("{}/{} Analyzing model '{}' with orifice status: {}".format(modelNameIndex+1, len(modelNames), modelName, orificeStatus

    # Copy dataframe and drop useless columns:
    LR_columns = df.columns.drop("DateTime")
    if orificeStatus=="NO":
        LR_columns = LR_columns.drop("OrificeSize")
    LR_df = df[LR_columns]
    LR_columns_validation = df_validation.columns.drop("DateTime")
    if orificeStatus=="NO":
        LR_columns_validation = LR_columns_validation.drop("OrificeSize")
    LR_df_validation = df_validation[LR_columns_validation]

    # Randomly shuffling the data (usually a good idea):
    np.random.seed(33)
    LR_df = LR_df.reindex(np.random.permutation(LR_df.index))
    LR_df.reset_index(inplace=True, drop=True)

    if orificeStatus=="WO":
        orifice = encode_text_index(LR_df, "OrificeSize")
        orificeValidation = encode_text_index(LR_df_validation, "OrificeSize")

    # Encode to a 2D matrix for training
    x,y = to_xy(LR_df, targetFeature)
    xValidation,yValidation = to_xy(LR_df_validation, targetFeature)

    # Split into train/test
    x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.25, random_state=45)

    # Create linear regression:
    regressor = sklearn.linear_model.LinearRegression()

    # Fit/train linear regression:
    regressor.fit(x_train,y_train)

    # Predict:
    pred = regressor.predict(x_test)
    predValidation = regressor.predict(xValidation)

    # Saving model description and its score to the main array:
    currentAnalysisResult = provideAnalysisResults(pred, y_test, predValidation, yValidation)
    modelSummary.append(currentAnalysisResult)
    print(currentAnalysisResult)

    # Saving the model to disk:
    modelFileName = ("{}\\{}{}.sav".format(folderForTheAnalysis, saveName, orificeStatus))
    pickle.dump(regressor, open(modelFileName, "wb"))

    plt.clf() # clearing all the plots from previous analysis
    # Plot the chart:
    expandedRegressionChart(pred, y_test, rollingWindowWidth=100)

    # Remove target feature from the array with all the column names:
    names = list(LR_df.columns.values)
    names.remove(targetFeature)

    # Plot regresssion coefficient importance:
    report_coef(names, regressor.coef_[0,:], regressor.intercept_)
    plt.show()
```

Fig. 6. Example of mathematical model implementation: Linear Regression

4. Machine Learning Simulation and Results Discussion

In order to properly establish acceptability threshold for model's results a benchmark method has been proposed. It consists of three metrics used to compare results of applied machine learning models: Mean Absolute Percentage Error (MAPE), Root-Mean-Square Error (RMSE) and Coefficient of Determination (R2). The three parameters have been chosen as they represent a basis for qualitative and quantitative comparison of the researched methods. MAPE metric allows valuable insight into the absolute magnitude of the error, while RMSE provides additional intuition into scaled residual size. R2 score represents a statistical measurement of model fit into predicted dataset. Target feature mean has been assumed as baseline prediction, resulting in RMSE of 0.83, MAPE of 9.24% and R2 score of 0 (per R2 score definition).

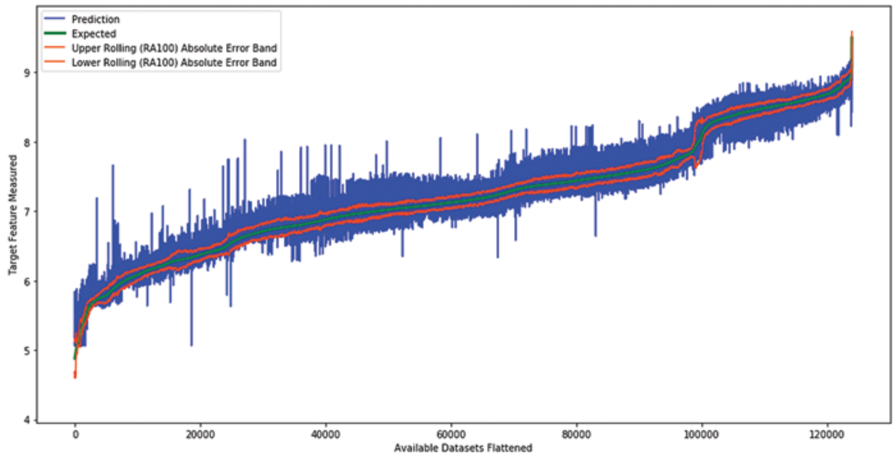


Fig. 7. Deep Learning model results. Due to large number of available datapoints (rows) Rolling Average of 100 samples (RA100) has been used to better visualize model results

Figure 7 shows an example of HP Recoup pressure parameter prediction postprocessing using Deep Learning model with 5-fold cross validation. Similar plots were created for each of the other 31 algorithms evaluated in this article. Figure 7 shows test set target feature range sorted ascending and its corresponding prediction using optimized Deep Learning model (as an example). The figure shows that while model error is generally consisted throughout entire model, the error increases significantly for certain feature inputs across entire target feature range. To prevent single set of measurements from triggering false negative alarms for the customer due to large model error it has been decided to smooth out the overall outcome using rolling average of the last 100 predicted values.

The result error has been shown on the chart in red and represents +/−MAPE band for the used rolling average window. As shown, it significantly reduces a single misprediction impact on overall maintenance system with a drawback of increase in system response time (up to 500 minutes of real time operations, depending on the parameter recording frequency). As overall system architecture is still being adjusted, the exact approach to filter out outlier predictions is not yet fully defined.

Altogether, 11 Scikit-learn machine learning models have been trained and used for comparison purposes followed by Keras (with Tensorflow backend) Artificial Neural Network (ANN) sequential dense models. Each model training process has been proceeded with extensive hyperparameter/parameter tuning on reduced dataset.

First, linear models, such as Linear Regression, Ridge Regression Model,Ridge Cross-Validation Regression Model, Kernel Ridge Regression, Lasso Regression, Lasso Cross-Validation and Elastic Net Model have been tested followed by Bayesian Ridge Regression Model (based on Bayes' theorem, that was applied to aviation gas turbines [4]). Then, more recently developed machine learning models, such as Support Vector Regression Model, Random Forest, and Gradient Boosting were analyzed. Finally, Artificial Neural Networks (ANN) with up to 3 hidden layers have been tested to allow comparison of established machine learning models to Deep Learning model.

The results of analysis are shown on Fig. 8 and 9 and summarized in Table 1. The most effective prediction has been achieved using Random Forest Regression model (0.018 RMSE, 0.122% MAPE and 0.9995 R2). Results acquired for the model, which included additional geometrical information (orifice size), were 0.01% higher (measuring the Coefficient of Determination, R2) than for the same model without additional data. Lower R2 results were achieved with Gradient Boosting Regression and simple Deep Learning model with 3 layers of Artificial Neural Networks (ANN). These models, however, achieved higher MAPE results (above 0.6% and 1.2% respectively). The first linear model that showed high R2 value was Support Vector Regression model achieving 0.155 RMSE, 1.258% MAPE and 0.965 R2 (additional geometrical data did not result in significant accuracy increase).

Random Forest Model has achieved highest performance using 80 estimators, while Gradient Boosted model shown lowest error rate with 245 estimators (and hinge at approximately 50 estimators).

As mentioned briefly in the previous chapter, in case of Artificial Neural Networks (ANN) models, training process consists of calculation of predefined loss function for given outputs and calculating

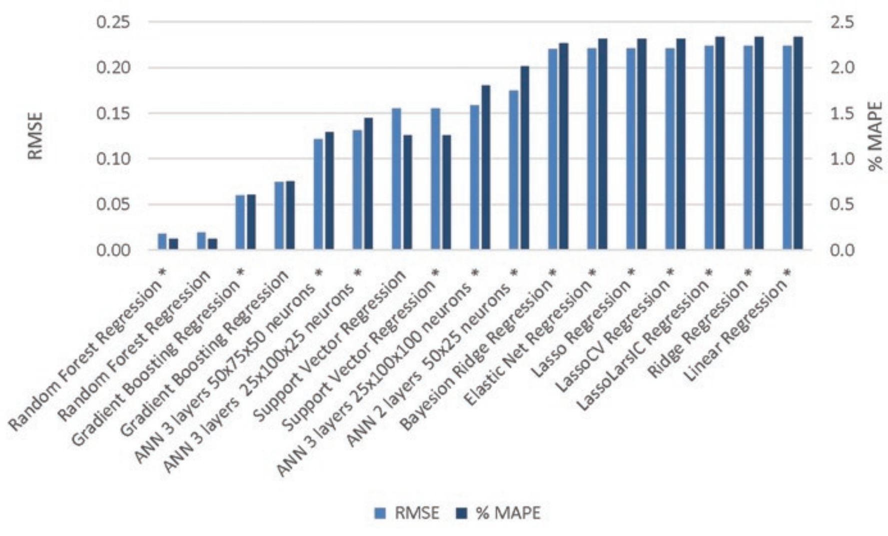


Fig. 8. The Root-Mean-Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) comparison for chosen Machine Learning models. Models with additional geometrical parameter (orifice size) are marked with asterix (*)

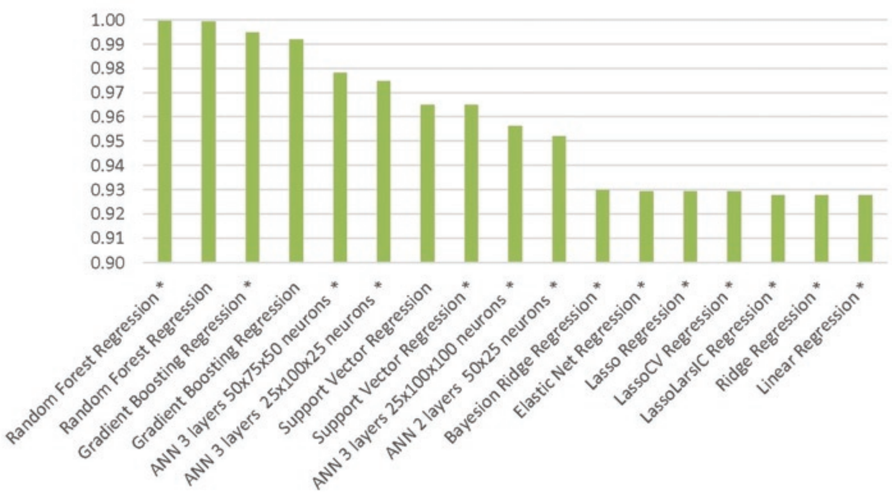


Fig. 9. Coefficient of Determination (R2) comparison for chosen Machine Learning models. Models with additional geometrical parameter (orifice size) are marked with asterix (*)

Table 1. Comparison of RMSE, MAPE and R2 for researched Machine Learning models. Models with additional geometrical parameter (orifice size) are marked with asterix (*)

	RMSE	% MAPE	R2 Score:
Random Forest Regression *	0,018	0,122	1,000
Random Forest Regression	0,019	0,125	0,999
Gradient Boosting Regression *	0,060	0,606	0,995
Gradient Boosting Regression	0,074	0,752	0,992
ANN 3 layers 50x75x50 neurons *	0,122	1,294	0,978
ANN 3 layers 25x100x25 neurons *	0,131	1,449	0,975
Support Vector Regression	0,156	1,259	0,965
Support Vector Regression *	0,156	1,259	0,965
ANN 3 layers 25x100x100 neurons *	0,159	1,803	0,956
ANN 2 layers 50x25 neurons *	0,175	2,013	0,952
Bayesion Ridge Regression *	0,220	2,270	0,930
Elastic Net Regression *	0,221	2,316	0,929
Lasso Regression *	0,221	2,315	0,929
LassoCV Regression *	0,221	2,315	0,929
LassoLarsIC Regression *	0,224	2,337	0,928
Ridge Regression *	0,224	2,335	0,928
Linear Regression *	0,224	2,336	0,928
Linear Regression	0,287	3,006	0,881
LassoLarsIC Regression	0,287	3,008	0,881
Ridge Regression	0,287	3,006	0,881
Bayesion Ridge Regression	0,287	2,998	0,881
Elastic Net Regression	0,288	3,030	0,880
Lasso Regression	0,288	3,037	0,880
LassoCV Regression	0,288	3,037	0,880
Kernel Ridge Regression	0,515	5,873	0,617
ANN 3 layers 25x25x100 neurons *	0,416	4,572	0,582
ANN 3 layers 50x100x100 neurons *	0,422	4,728	0,581
ANN 3 layers 50x25x75 neurons *	0,479	5,478	0,543
Kernel Ridge Regression *	0,646	7,606	0,398
RidgeCV Regression *	0,779	8,814	0,125
RidgeCV Regression	1,865	21,447	-4,023

partial derivatives across different layers for all weights (parameters representing connection strength between nodes) and biases (parameters used to adjust the model output to fit best data set). Recent years have shown increase in model efficiency through implementing new design of activation functions (including ReLU, used in this analysis), initialization methods, regularization methods, learning rate adjustment (including ADAM, also used in this thesis) and input sampling (SGD) while GPU/TPU based increase in computation power significant reduction in training time. While computer disc memory size consistently increases, introduction of Batch Gradient Descent allows training on theoretically infinite size dataset by subsampling it into smaller pieces capable of fitting into memory. As such Deep Learning allows user to operate and generalize much larger problems and enables end-to-end solution generation for any machine learning problem. This contrasts with machine learning models such as SVC, which are limited by memory size and can only operate on user predefined features.

Different size of 3 hidden layers have been tested within a range of 25-100 and an increment of 25 for each layer. Optimum performance has been obtained for a network with 25 nodes in the layer 1, 25 nodes in layer 2 and 100 nodes in layer 3.

Initially team considered utilization of additional categorical feature, HP Recoup pressure orifice size. The analysis has shown that it did not play significant role in case of complex models' application, such as Random Forest or Gradient Boosting. At the same time most of the linear models showed significantly higher R2 results for models with additional data. For example, Lasso Regression had 5.6% R2 score higher while Kernel Ridge Regression (generally one of the least performing models) achieved 54.9% difference.

Presented approach has been validated on gas turbines with operational range very similar to the training data, which partially skews presented results up. Expansion of available dataset for additional gas turbines with further hyperparameter tuning should increase model generalization capability while maintaining its high accuracy. Furthermore, analysis considered multivariate approach for single output, which results in time series information loss and relatively high point-to-point model results variation discovered during data postprocessing. Variation could be handled by simple methods like rolling average. Further increase in model accuracy can be achieved by extracting data sequence information. For example, Recurrent Neural Network could be used on top of preliminary model (ML/DL) results to reduce target feature variation and increase accuracy.

5. Conclusions

The presented study shows comparison of the gas turbine engine HP Recoup pressure prediction. It may be used for gas turbine predictive maintenance planning, potentially allowing gas turbine cost model improvement and optimization, as shown by Deloux et al. [9]. The applicability of numerical machine learning (ML) based prediction models for gas turbine operating parameters prediction has been demonstrated. To specify acceptable model threshold a basic (target data average) benchmark has been proposed and used. In order to allow quantitative comparison between different models 31 machine learning algorithms have been tested, including Artificial Neural Networks, random forests, boosted random forest and SVC. Best results were obtained for random forest regression due to its quick generalization capability enabling ensemble solution with relatively low computational power. Gradient boosting methods also have shown high accuracy due to residua minimization approach utilized in algorithm design. Artificial Neural Networks or deep learning methods were also shown application potential, showing high accuracy results with only 3 hidden layers. While computational requirement for deep learning hyperparameter tuning are significantly higher than those of random forest regressor [15], ANN model can be easily tuned and adjusted for other, similar problems (transfer learning), while majority of machine learning algorithms must be completely retrained for new purposes.

Moreover, data tests showed that additional geometrical data (such as orifice size, available for chosen gas turbines in researched dataset) is not always crucial to improve prediction quality, although it improves overall accuracy of the model and should be used if available, if only to check for model overfitting. Relatively simple numerical model result's comparison leads to most appropriate model, that guarantees high accuracy. Currently Artificial Neural Networks (ANN) architectures could also offer valuable insight such as prediction confidence, which would be critical for applications such as predictive maintenance.

Lastly, the developed methodology is applicable to any of the gas turbine parameter, when reference physics-based models and dataset from sufficiently large fleet are available to validate the accuracy of the data-driven algorithms developed. Achieved results showed that high accuracy may be obtained using the same input data, but with different machine learning algorithms after extensive hyperparameter tuning.

Acknowledgement

The work is based on Baker Hughes industrial data and has been developed as part of GE Challenge funded by General Electric Company Polska Sp. z o.o [19].

References

1. Augustyn S. Energy Model of Change in Technical Condition of Aircraft Power Plants and Space Propulsion Systems. Aviation Advances & Maintenance 2017; 40(2), https://doi.org/10.1515/afit-2017-0009.
2. Badeer G. H. Engine Thrust Bearing Condition Monitoring Method. US Patent US no. 6,637,932 B2, GE Schenectady, NY, US, Oct. 23, 2003.
3. Badeer G. H. GE Aeroderivative Gas Turbines - Design and Operating Features. GE Power Systems GER-3695E (10/00) available online: https://www.ge.com/content/dam/gepower-pgdp/global/en_US/documents/technical/ger/ger-3695e-ge-aero-gas-turbine-design-op-features.pdf
4. Batalha E. Aircraft Engines Maintenance Costs and Reliability. An Appraisal of the Decision Process to Remove an Engine for a Shop Visit Aiming at Minimum Maintenance Unit Cost 2012.
5. BHGE aeroderivative gas turbines portfolio available online at https://www.bhge.com/aeroderivative-gas-turbines
6. Carlevaro F, Cioncolini S, Sepe M, Parrella I, Escobedo E, Allegorico C, De Stefanis L, Mastroianni M. Use of operating parameters, digital replicas and models for condition monitoring and improved equipment health. ASME Turbo Expo 2018, https://doi.org/10.1115/GT2018-76849.
7. Chmielewski M, Fulara S, Gieras M. Numerical Prediction of GTD-350 Turboshaft Engine Combustor Deterioration. Journal of KONES 2017; 24(2): 47-58.
8. Cyrus B Meher-Homji C. B, Yates D, Weyermann H. P. Aeroderivative Gas Turbine Drivers for The ConocoPhillips Optimized CascadeSM LNG Process - World's First Application and Future Potential. 15th International Conference & Exhibition on Liquefied Natural Gas 2007 available online: http://www.ivt.ntnu.no/ept/fag/tep4215/innhold/LNG%20Conferences/2007/fscommand/PS2_6_Meher_Homji_s.pdf
9. Deloux E, Castanier B, Berenguer C. Predictive maintenance policy for a gradually deteriorating system subject to stress. Reliability Engineering and System Safety 2009; 94: 418- 431, https://doi.org/10.1016/j.ress.2008.04.002.
10. Demgne J, Mercier S, Lair W, Lonchampt J. Modelling and numerical assessment of a maintenance strategy with stock through Piecewise Deterministic Markov Processes and Quasi Monte Carlo methods. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2017; 231(4): 429-445, https://doi.org/10.1177/1748006X17712121.
11. Fulara S, Chmielewski M, Gieras M. Experimental research of the small gas turbine with variable area nozzle. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering 2019; 233(15): 5650-5659, https://doi.org/10.1177/0954410019853977.
12. 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.
13. Garcia Nieto P.J, Garcia-Gonzales E, Sanches Lasheras F, de Cos Juez F.J. Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability. Reliability Engineering & System Safety 2015; 138: 219-231, https://doi.org/10.1016/j.ress.2015.02.001.
14. GE Power & Water Distributed Power. LM2500+ G4 BaseGas Turbine Fact Sheet available online: https://www.ge.com/content/dam/gepower-pgdp/global/en_US/documents/product/lm2500-plus-g4-fact-sheet.pdf accessed on September 2018
15. Heaton J. Automated Feature Engineering for Deep Neural Networks with Genetic Programming. College of Engineering and Computing Nova Southeastern University 2017.
16. Herbert L. Designing for Reliability, Maintainability, and Sustainability (RM&S) in Military Jet Fighter Aircraft Engines, Massachusetts Institute of Technology 2002.
17. Huang B, Wang Z. The Role of Data Prefiltering For Integrated Identification and Model Predictive Control. IFAC Proceedings Volumes 1999; 32(2): 6751-6756, https://doi.org/10.1016/S1474-6670(17)57153-0.

18. Huang H.Z, Tong X, Zuo M. J. Posbist fault tree analysis of coherent systems. Reliability Engineering & System Safety 2004; 84(2): 141-148, https://doi.org/10.1016/j.ress.2003.11.002.
19. Institute of Aviation. GE Challenge available online: https://ilot.edu.pl/tag/ge-challenge accessed on August 2018
20. Jankowski A, Kowalski M. Creating Mechanisms of Toxic Substances Emission of Combustion Engines. Journal of KONBiN 2015; 36(1): 33-42, https://doi.org/10.1515/jok-2015-0054.
21. Kopytov E, Labendik V, Yunusov S, Tarasov A. Managing and Control of Aircraft Power Using Artificial Neural Networks. Proceeding of the 7th International Conference "Reliability and Statistics in Transportation and Communication 2007.
22. Kozik P, Sęp J. Aircraft Engine overhaul demand forecasting using ANN. Management and Production Engineering Review 2012; 3(2): 21-26.
23. Kumar U. D, Crocker J, Knezevic J. Evolutionary Maintenance for Aircraft Engines. Proceedings Annual Reliability and Maintainability Symposium 1999: 62-68.
24. Liu D, Zhang H, Polycarpou M, Alippi C, He H. Elman-Style Process Neural Network with Application to Aircraft Engine Health Condition Monitoring. Advances in Neural Networks. Proceedings of the 8th International Symposium on Neural Networks 2011.
25. Louppe G. Understanding Random Forests: From Theory to Practice. ArXiv Preprint 2014; ArXiv:1407.7502: 35.
26. Lu J-M, Innal F, Wu X-Y, Liu Y, Lundteigen M. A. Two-terminal Reliability Analysis for Multi-Phase Communication Networks. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(3): 418-427, https://doi.org/10.17531/ein.2016.3.14.
27. 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.
28. Michelassi V, Allegorico C, Cioncolini S, Graziano A, Tognarelli L, Sepe M. Machine learning in gas turbines, from component design to asset management. ASME Journal: Global Gas Turbine News 2018; 140(09): 54-55, https://doi.org/10.1115/1.2018-SEP5.
29. Mingazov B. G, Korobitsin N. A. Use of Probability Indices for Assessment of Gas Turbine Power Station Reliability under Commercial Operation Conditions. Russian Aeronautics 2009; 53(2): 226-229, https://doi.org/10.3103/S1068799810020170.
30. Morris A.S, Langari R. Measurement and Instrumentation: Theory and Application 2016, https://doi.org/10.1016/B978-0-12-800884-3.00017-4.
31. Pawełczyk M, Bibik P. Wykorzystanie nowoczesnych narzędzi inżynierskich w projektowaniu bezzałogowego wiropłata czterowirnikowego. Materiały IX Krajowego Forum Wiropłatowego. Instytut Lotnictwa 2013, https://doi.org/10.5604/05096669.1106572.
32. Simon D.L. An Integrated Architecture for On-Board Aircraft Engine Performance Trend Monitoring and Gas Path Fault Diagnostics. 57th Joint Army-Navy-NASA-Air Force (JANNAF) Propulsion Meeting sponsored by the JANNAF Interagency Propulsion Committee 2010.
33. Soumitra P, Kapoor K, Jasani D, Dudhwewala R, Gowda V. B, Gopalakrishnan Nair T.R. Application of Artificial Neural Networks in Aircraft Maintenance. Repair and Overhaul Solutions, Analysis and Manufacturing Technologies 2008.
34. Tibshirani R, Friedman J, Hastie T. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software. 2010; 33(1), https://doi.org/10.18637/jss.v033.i01.
35. Verma M, Kumar A. A novel general approach to evaluating the reliability of gas turbine system. Engineering Applications of Artificial Intelligence 2014; 28: 13-21, https://doi.org/10.1016/j.engappai.2013.10.001.
36. Walsh P, Fletcher P. Gas Turbine Performance. Blackwell Science 2004, https://doi.org/10.1002/9780470774533.
37. Zadeh L.A. Fuzzy sets. Information and Control 1965; 8(3): 338-353, https://doi.org/10.1016/S0019-9958(65)90241-X.
38. Zhang H, Keerthi S, Mahajan D, Dhillon I.S, Hsieh C.J. Gradient Boosted Decision Trees for High Dimensional Sparse Output. In ICML; 2017: 3182-3190.

Maciej PAWEŁCZYK
Szymon FULARA
Łukasiewicz Research Network – Institute of Aviation
al. Krakowska 110/114
02-256 Warsaw, Poland

Marzia SEPE
Alessandro DE LUCA
Baker Hughes
Via Felice Matteucci 2
50127 Firenze, Italy

Maciej BADORA
Baker Hughes
al. Krakowska 110/114
02-256 Warsaw, Poland

E-mails: maciej.pawelczyk@ge.com, szymon.fulara@ge.com, marzia.sepe@bakerhughes.com, alessandro.deluca@bakerhughes.com, maciej1.badora@bakerhughes.com

Article citation info:
ÖZCAN E, DANIŞAN T, YUMUŞAK R, EREN T. An artificial neural network model supported with multi criteria decision making approaches for maintenance planning in hydroelectric power plants. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 400–418, http://dx.doi.org/10.17531/ein.2020.3.3.

Evrencan ÖZCAN
Tuğba DANIŞAN
Rabia YUMUŞAK
Tamer EREN

AN ARTIFICIAL NEURAL NETWORK MODEL SUPPORTED WITH MULTI CRITERIA DECISION MAKING APPROACHES FOR MAINTENANCE PLANNING IN HYDROELECTRIC POWER PLANTS

PLANOWANIE UTRZYMANIA RUCHU W ELEKTROWNIACH WODNYCH W OPARCIU O MODEL SZTUCZNEJ SIECI NEURONOWEJ WSPARTY WIELOKRYTERIALNYMI METODAMI PODEJMOWANIA DECYZJI

Power plants are the large-scale production facilities with the main purpose of realizing uninterrupted, reliable, efficient, economic and environmentally friendly energy generation. Maintenance is one of the critical factors in achieving these comprehensive goals, which are called as sustainable energy supply. The maintenance processes carried out in order to ensure sustainable energy supply in the power plants should be managed due to the costs arising from time requirement, the use of material and labor, and the loss of generation. In this respect, it is critical that the fault dates are forecasted, and maintenance is performed without failure in power plants consisting of thousands of equipment. In this context in this study, the maintenance planning problem for equipment with high criticality level is handled in one of the large-scale hydroelectric power plants that meet the quintile of Turkey's energy demand as of the end of 2018. In the first stage, the evaluation criteria determined by the power plant experts are weighted by the Analytical Hierarchy Process (AHP), which is an accepted method in the literature, in order to determine the criticality levels of the equipment in terms of power plant at the next stage. In order to obtain the final priority ranking of the equipment in terms of power plant within the scope of these weights, Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is used because of its advantages compared to other outranking algorithms. As a result of this solution, for the 14 main equipment groups with the highest criticality level determined on the basis of the power plant, periods between two breakdowns are estimated, and maintenance planning is performed based on these periods. In the estimation phase, an artificial neural network (ANN) model has been established by using 11-years fault data for selected equipment groups and the probable fault dates are estimated by considering a production facility as a system without considering the sector for the first time in the literature. With the plan including the maintenance activities that will be carried out before the determined breakdown dates, increasing the generation efficiency, extending the economic life of the power plant, minimizing the generation costs, maximizing the plant availability rate and maximizing profit are aimed. The maintenance plan is implemented for 2 years in the power plant and the unit shutdowns resulting from the selected equipment groups are not met and the mentioned goals are reached.

Keywords: *Artificial neural networks, hydroelectric power plants, failure period estimation, maintenance planning, AHP, TOPSIS.*

Elektrownie to zakłady produkcyjne o dużej skali, których głównym celem jest nieprzerwane, niezawodne, wydajne, rentowne oraz przyjazne dla środowiska wytwarzanie energii. Utrzymanie ruchu stanowi jeden z kluczowych czynników pozwalających na osiągnięcie tych szeroko zakrojonych celów, które określa się wspólnym mianem zrównoważonych dostaw energii. W elektrowniach, procesami utrzymania ruchu, realizowanymi w celu zapewnienia zrównoważonych dostaw energii, zarządza się z uwzględnieniem kosztów związanych z wymogami czasowymi, kosztów materiałów i robocizny oraz strat wytwarzania energii. Ponieważ elektrownie wykorzystują tysiące różnych urządzeń, niezwykle ważne jest prognozowanie dat wystąpienia uszkodzeń oraz zapewnienie bezawaryjnego utrzymania ruchu. W przedstawionych badaniach, rozważano problem planowania utrzymania ruchu sprzętu o wysokim poziomie krytyczności na przykładzie jednej z dużych elektrowni wodnych, która na koniec 2018 r. pokrywała jedną piątą zapotrzebowania Turcji na energię elektryczną. W pierwszym etapie badań, kryteria oceny określone przez ekspertów zatrudnionych w elektrowni ważono za pomocą powszechnie stosowanej w literaturze metody procesu hierarchii analitycznej (AHP) w celu ustalenia poziomów krytyczności poszczególnych elementów wyposażenia elektrowni. Aby opracować ostateczny ranking priorytetowości elementów wyposażenia elektrowni na podstawie określonych wcześniej wag, zastosowano technikę TOPSIS, która polega na porządkowaniu preferencji na podstawie podobieństwa do idealnego rozwiązania. Techniki tej użyto ze względu na jej zalety, których nie mają inne algorytmy oparte na relacji przewyższania (ang. outranking algorithms). Na podstawie wyników otrzymanych dla 14 głównych grup urządzeń o najwyższym poziomie krytyczności, określonym na podstawie danych pochodzących z elektrowni, oszacowano czasy pomiędzy dwiema awariami, a na ich podstawie zaplanowano działania konserwacyjne. W fazie szacowania, opracowano model sztucznej sieci neuronowej (ANN) w oparciu o dane o uszkodzeniach, które wystąpiły w ostatnich 11 latach działania elektrowni, dla wybranych grup urządzeń. Przewidywane daty wystąpienia uszkodzeń szacowano, po raz pierwszy w literaturze, biorąc pod uwagę zakład produkcyjny jako system, bez uwzględnienia sektora *produkcyjnego. Plan obejmuje działania konserwacyjne, które mają być przeprowadzone przed przewidywanymi datami awarii, w celu zwiększenia wydajności wytwarzania energii, przedłużenia żywotności elektrowni, minimalizacji kosztów wytwarzania energii, maksymalizacji wskaźnika dostępności elektrowni oraz maksymalizacji zysków. Opracowany plan konserwacji wdrażano w omawianej elektrowni przez 2 lata. W tym okresie nie odnotowano przerw w pracy jednostek wytwórczych spowodowanych awarią rozważanych grup urządzeń, co oznacza, że wspomniane cele zostały osiągnięte.*

Słowa kluczowe: *sztuczne sieci neuronowe, elektrownie wodne, szacowanie czasu między uszkodzeniami, planowanie utrzymania ruchu, AHP, TOPSIS.*

1. Introduction

In today's conditions where competition is increasing, enterprises need to increase their profitability in order to ensure their continuity in their activities or to include plans aimed at minimizing the total cost resulting from these activities. These plans are of great importance in order to manage the operational activities that can be classified under the titles of production, personnel, material and maintenance in a systematic and effective manner. Especially in order to realize these targets, it is critical that the machinery, equipment and devices in the production facilities perform the expected functions in a timely, uninterrupted, high quality and reliable manner. In this respect, planning of the maintenance activities contributes greatly to the effective management of the other three main processes (production, personnel and material), while systematic management of maintenance planning activities plays an important role in achieving improvements in operational efficiency.

Besides contributions provided by the mentioned maintenance planning activities, it is necessary to optimize all the costs in accordance with the operating conditions as maintenance costs can reach 15-70% of the various production costs varying according to the type of operation [11]. According to the most commonly used method, maintenance cost is composed of labor, spare parts and service costs which are spent for maintenance [19]. In case of stoppages caused by malfunction or any maintenance application in the enterprise, the losses for each time period in which production cannot be realized should be considered as cost and included in the management process. The most important point to be considered here is the monitoring, inspection and follow-up taking them under control of maintenance performance parameters such as the mean time between failures (MTBF), failure stop rate (FSR) and fault repair time (FRT). Recording the data in which the maintenance performance indicators are mentioned in the maintenance practices realized by the enterprises is one of the important factors that will contribute to this process.

At the same time, hydroelectric power plants are critical with about one fifth share in Turkey's energy mix [21]. Therefore, a large-scale hydroelectric power plant in Turkey is selected as application field, and this plant has 6,111 equipment. It is not possible to carry out maintenance so many equipment in terms of continuity of generation and minimization of the costs. In fact, every equipment in the power plant does not directly affect sustainable energy generation. In other words, the level of impact of equipment on sustainable energy supply can be expressed as the level of risk (criticality level) of equipment in terms of power plant. For this reason, prioritization of maintenance activities according to criticality levels of equipment would not be deviated the power plant from its sustainable energy supply goal but would serve this comprehensive purpose. Furthermore, considering that the purpose of the maintenance is to extend the time of the fault-free operation of the equipment, it will be possible for an equipment to reach this goal in the most appropriate way (especially in terms of cost efficiency) by the maintenance before the possible downtime. From this point of view in this study, a new maintenance planning methodology is proposed for the efficient maintenance management and hence an efficient power plant management, which includes the maintenance performance indicators in a big-scale hydroelectric power plant with the aim of achieving continuous, reliable, efficient, economic and environmentally friendly electricity generation [58].

First, 9 criteria which affect the criticality level of the equipment in the power plant are weighted by AHP, and the obtained weights are used in TOPSIS algorithm for calculating the risk levels of equipment in terms of sustainable energy supply in the power plant. It is determined that 14 equipment groups have the maximum risk level in this ranking obtained for 6,111 equipment. This calculation is consistent with real life in terms of threatening the uninterrupted, reliable, efficient, economic and environmentally friendly power generation when 14 equipment groups fail. Then, an ANN model is established by using 11-years fault data including maintenance performance indicators for selected equipment groups and possible fault dates are estimated, and maintenance schedule is based on the estimated date of the failure for each equipment for the first time in the literature. As a result of the implementation of this schedule in the power plant for 2 years, the generation stoppages resulting from the lack of maintenance in the selected equipment groups are reduced by 100%. In addition, increasing the generation efficiency, extending the economic life of the power plant, minimization of generation costs, maximization of power plant availability and profit maximization are achieved.

In the second section of the study, the studies in literature about the problem in a broad perspective are included. In the third section, the methods used in the study are given with use case and the application phase is described in section 4. The study is completed with section 5 where the results and recommendations are presented.

2. Related literature

Multi-Criteria Decision Making (MCDM) is an approach that makes decision-making more effective when there are often conflicting and/or related, qualitative and quantitative criteria [31]. The fact that the problem parameters are qualitative or quantitative in the decision-making process and that these parameters should be evaluated together make the MCDM a practical and comprehensive evaluation strategy. For this reason, in many studies in the literature, many MCDM methods, such as AHP [33,39,77], Analytic Network Process (ANP) [73], TOPSIS [10], ELimination Et Choix Traduisant la REalité (ELECTRE) [24], Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) [9] and Vlsekriterijumska Optimizacija I KOmpromisno Resenje (VIKOR) [26] are used in different problem areas such as location [36], project [44,66], staff [4], supply chain [81] and strategy selection [15], and health applications [47]. Furthermore, energy related decision-making problems have also intrinsically multiple criteria structures and therefore, analytical approaches for effective solutions for these problems are needed. In this context and within the frame of the advantages of MCDM stated above, this approach provides effective results in solving many problems related with energy. Especially, the reviews performed by Mardani et al. [50] and Kumar et al. [40] have proved the importance of the MCDM to be very important for the problems in energy related studies. Another point that is noteworthy is that the usage of method combinations composed of one more than techniques under MCDM approach for problem solving have enabled us to receive effective results. For example, Zyoud and Fuchs-Hanusch [92] have proved that the studies in which AHP and TOPSIS which are also used in this study are integrated, are more effective than the separate usage of the methods. The studies performed by Özcan et al. [58] and Kumar et al. [41] in the hydroelectric power plants, and Sindhu et al. [72] in the solar power plant are the important studies that integrate these two methods. In this study, the solution is started with AHP method due to the subjectivity reduction, ease of application, widespread use and flexibility of integration with linear programming, fuzzy logic and especially sorting algorithms [39,77,82]. In the second phase, the TOPSIS method is preferred because of the easy and effective way to perform the alternative ranking [10]. Another method used in the study, the ANN is used together with these MCDM methods in a few studies in the literature [8,42,76,84]. Today, ANN frequently encountered among artificial intelligence methods such as fuzzy logic, genetic algorithms etc. is used for solving the many types of problems about classification [46,90], diagnostics [3,91], scheduling [45] and prediction [2,80,89] in the wide sectoral range such as atmospheric sciences [23], transportation [20], finance [87], health [69] and energy [37].

Recently, prediction studies focused on ANN [70,89] and obtaining effective results is one of the factors that highlight this method. In addition, traditional prediction methods have been sufficient due to the long-term linear problems in the prediction studies in previous years [13,59]. Nowadays it is not reasonable to assume that the problems are linear because of the dynamic and variable structure of real-life problems [25]. Therefore, these methods are incomplete in some respects in nonlinear problems. In contrast to the traditional prediction models, ANN models learn the structure of the problem from historical data on the problem and captures a fine functional relationship between the parameters even if the problem is difficult. This is the most important feature of ANN models, which shows their superiority compared to traditional ones in nonlinear problems. Therefore, the use of ANN models is more effective for problems that are especially difficult to solve where sufficient data or observation can be obtained [18,62,86]. In addition, it is necessary to transfer large amounts of data from the system examined in deep learning applications [61]. However, there are 285 faults in the hydroelectric power plants, where the problem is addressed, in the 11-year period. While this size of data set is insufficient for deep learning methods, it is sufficient for ANN method. Therefore, ANN method was preferred in this study instead of both traditional statistical methods and deep learning methods. Energy-related problems are often complex, and therefore the ANN method is used effectively in forecasting problems related to energy. The literature reviews performed by Suganthi and Samuel [74] on electricity demand, Weron [85] on electricity prices, and Wang and Srinivasan [83] on energy usage, evidence by the fact that the ANN is frequently used in the energy related studies.

ANN can learn the process based on historical data and obtain highly accurate estimation results even if changes occur in the process [55]. In addition, the applications of artificial intelligence, including the ANN, in power plants aim to minimize human intervention in processes. In other words, these practices aim to reduce the dependence on staff in all process management, including maintenance planning, because of the threats that they may cause [51]. As a result of these objectives and the advantages mentioned above, ANN has taken place in the maintenance planning problem in the energy sector. Examples of maintenance planning literature in the energy sector are summarized below:

Messai et al. [53] performed a system design that forecasts the temperature of the fuel rod temperature sensor in the nuclear reactor core by a Bayesian Network, which is a special ANN model. Ayodeji et al. [6] assessed the performance of the simulation with two ANN models predicting the location and size of the fault before the malfunction occurred in the operator support system. As with these studies, ANN models not only use environmental variables such as outdoor temperatures or radiation, but also the condition variables of the equipment as internal temperature for different operating times. In this way, faults can be detected before affecting production and a quantitative risk measurement can be realized. From this point of view, Polo et al. [57] made the malfunction mode and energy production estimation of critical equipment for photovoltaic power plant with ANN.

Most of the maintenance planning studies carried out with ANN in the renewable energy field are carried out in wind power plants which are advantageous due to simple production structure according to other renewable energy plants. The study of Schlechtingen and Santos [67], which deals with offshore wind turbines, is based on the Supervisory Control and Data Acquisition (SCADA) data of 10 wind turbines of the same type with 2 MW installed capacity. By using these data, two different ANN models are constructed for fault estimation and the results are compared. Faults are evaluated as first and second faults on a day-by-day basis and temperature values are estimated. Based on these estimations, the fault is detected at the earliest 50 days ago. Kusiak and Li [43], the other researchers working on the estimation of faults, estimated the failures in three phases based on data from SCADA of 4 wind turbines. In this study, the first stage is to determine the presence of the fault, the second stage is the estimate of the severity of the fault and the final stage is the specific fault estimation. Three-phase fault estimation is modeled by using Neural Network (NN), Neural Network Ensemble (NN Ensemble), Boosting Tree Algorithm (BTA) and Support Vector Machine (SVM) methods and it was able to find the problem 5-60 minutes before failure occurred. Another study on wind turbines is performed by Chen et al. [16]. In this study, Adaptive-Network Based Fuzzy Inference Systems (ANFIS) which is a smart approach to classify faults in wind turbines is used unlike fault estimation. There are other studies using the ANFIS method. The second study of Schlechtingen and Santos [68] is also an example to the ANFIS usage for maintenance planning. In this study, a system for monitoring wind turbines with ANFIS which is a combination of ANN and fuzzy logic analysis, is proposed according to SCADA data. In the continuation of the study, some sample applications of the system such as hydraulic failure, cooling system failure, anemometer failure and turbine control device failures are shown. Sun et al. [75] presented a generalized ANN model to estimate deviations of parameter data such as rotor speed, output power and component temperature collected in SCADA. Instead of testing the model on two states and using a single estimation model, multiple prediction models which are trained with different types of sample data are integrated to determine the deviations of the status parameters of the wind turbine. It is shown that the proposed method is more effective than the traditional single model-based method in the definition of turbine deviation values. In another study about wind turbines, Bi et al. [12] developed a new system that gives an alarm 13-20 hours before the current system by using SCADA data for generator with ANN and ANFIS integration. Bangalore and Patriksson [7] carried out an exemplary study of hybrid models on wind turbines. In this study, implementation of ANN based status monitoring method is presented for a wind turbine and the system has been able to detect the fault 2 months in advance. They also proposed a mathematical optimization model for preventive maintenance scheduling. Finally, Lu et al. [49] predicted the economic life of rotor, transmission and generator equipment in the wind turbine with ANN.

When the studies are examined in the literature, the ANN method is frequently used in fault prediction studies without considering the sector. The review on fracture mechanics using artificial intelligence methods by Nasiri et al. [56] classified the studies under the titles of fault mode and mechanism identification, damage and fault detection and diagnosis, error and error detection, the diagnosis and mechanical fracture parameters. As a result of the review, they confirmed that the ANN method is used in 46% of the studies examined. This finding proves that the ANN method is effectively used in the failure estimation problems.

In this study, the time estimation between the two faults (in other words, mean time between failures-MTBF) is discussed, and in the literature, studies on maintenance planning based on fault estimations based on maintenance performance parameters and evaluating the system in the light of these parameters come to the fore. For example, in order to investigate system reliability, Komal and Sharma [38] applied three techniques based on ANN and genetic algorithm for the washing system in a paper factory, Jiang et al. [34] modeled the effect of climate and environmental conditions in wind turbines using the regression method, and Vedachalam and Ramadass [78] carried out an exemplary design for dynamic command system. For the prediction of MTBF, which is effective in maintenance planning [34], Chen et al. [17] performed the MTBF estimation for a CNC machine using the DGM model because MTBF is an important parameter in the reliability of complex equipment. Braglia et al., presented a multivariate statistical approach that supports the classification of mechanical components in terms of MTBF [14]. They proposed a new approach to differentiate the operating parameters due to the differences in MTBF, and the determination of MTBF of mechanical parts, depending on the specific operating conditions. Unlike other authors, Jones et al. [35], Adoghe et al. [1] and Illias et al. [32] performed their studies by conducting investigations on malfunctions. Jones et al. [35] calculated the failure rate of the system with Bayesian Network for maintenance planning in a manufacturing industry and used 1 / MTBF ratio in this calculation. Adoghe et al. [1] performed the statistical analysis of the failure data using the serial correlation coefficient and Laplace test methods, and thus, they provided the selection of critical components having the highest risk index in terms of system reliability, and they presented an effective maintenance planning program to address these critical components. Thus, they proposed a reliability- centered maintenance methodology based on statistical analysis for an electrical distribution system. Illias et al. [32] used dissolved gas analysis with a combination of ANN and three PSO techniques to estimate the initial failure of the transformer. In the ANN model, they used 100 input data from 6 gas types, and considered the fault data, thermal error, low density and high- density parameters as output data. They categorized these data as training, verification, and test sets. Apart from these studies, Liu et al. [48] used the MTBF for the reliability of the CNC grinder and Yang et al. [88] used the maintenance failure repair time parameter.

Following the above explanations of the studies using ANN about maintenance planning, the differences of this study from other studies in the literature are as follows:

- In the literature, different artificial intelligence methods are used for maintenance planning problem, and the fault can be detected at the earliest 2 months ago. This study is based on the MTBF parameter, which presents a different philosophy between fault estimation studies. In this study, it is detected the faults earlier than other studies in the literature, and the estimation is made in periods between 201 and 1461 days. Predicting the faults in such an early stage proves the applicability of the model developed in real life problems, considering the length of the preliminary preparation and implementation process required for maintenance.
- In this study, maintenance planning is performed based on MTBF estimation. Similar studies in terms of MTBF estimation in maintenance planning are often carried out for wind turbines. In these studies, a few equipment (such as turbine rotor, transmission, generator or generally turbine for other power plants) is considered because these machines have not complex structure. However, hydroelectric power plants are the most complex one among the all renewable energy power plants. In this respect, this is the first study on maintenance planning based on the MTBF parameter in the literature, which deals with complex hydroelectric power plants in a system.
- Furthermore, hydroelectric power plants are the most mature renewable energy technology and therefore, these big-scale and complex plants affect the countries' energy mix more than the wind turbines (e.g. the shares of hydro and wind in total electricity generation in Turkey as of the end of 2018 are 20% and 6.3% respectively [21]). Accordingly, hydroelectric power plants are the most critical renewable resource for sustainable energy supply in the world. Therefore, it can be said that the effect of maintenance at these facilities on the sustainable energy supply is much higher than the wind turbines.
- Another fact that makes this study stand out is the use of AHP-TOPSIS-ANN method integration in terms of increasing the level of analyticalness in this problem field, which together with all these features make the study different. It is thought that this study will contribute significantly to the literature due to all these differences.

3. Methods

3.1. AHP

The AHP method developed by Saaty is used as a singular or supportive method in many decision-making problems and its popularity is increasing day by day. This method allows people to define priorities between criteria and alternatives in the decision-making process, together with qualitative and quantitative judgments [79].

The following are the implementation steps of AHP [65]:

Step 1: Determination of goal, criteria, sub-criteria, alternatives and hierarchical structure

This phase includes the aim of the decision maker, the criteria affecting this goal, and the determination of the relationships between them through the addition of alternatives (Figure 1).

Step 2: Performing the pairwise comparison for criteria and alternatives for each criterion

It is carried out by experts by comparing all criteria and alternatives according to their severity. At this stage, the 1-9 preference scale, which is developed by Saaty and given in Table 1 is used.

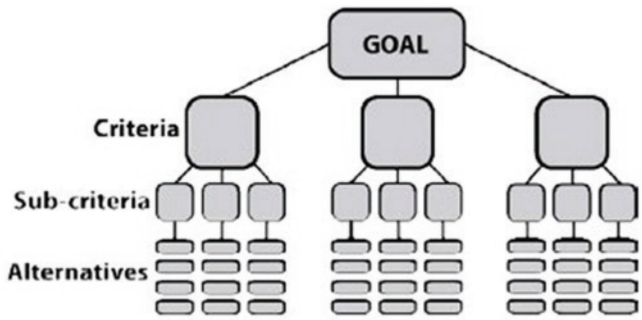


Fig. 1. Hierarchical structure [65]

Step 3: Calculation of priority vector

The vector weights (w) are calculated using the pairwise comparison matrix, normalization of $A.w = \lambda_{max}.w$, and the following equation:

Table 1. Saaty's preference scale [65]

Importance Values	Value Definitions
1	Equal importance of both factors
3	Factor 1 is more important than factor 2
5	Factor 1 is much more important than factor 2
7	Factor 1 has a very strong importance compared to factor 2
9	Factor 1 has an absolute superior importance to factor 2
2, 4, 6, 8	Intermediate values – when compromise is needed

$$w_i = \sum_{i=1}^{n} b_{ij} / n \tag{1}$$

Step 4: Calculation and control of the consistency ratio (CR)

CR is calculated as the result of the ratio of the consistency index (*CI*) to the random consistency index (*RI* - Table 2) (Eq.3). Eq.2 is used to calculate the *CI*:

$$CI = \left(\lambda_{max} - n\right) / \left(n-1\right) \tag{2)0$$

Table 2. RI values for different n values [65]

n	1	2	3	4	5	6	8	9	10	11	12	13
RI	0	0	0,58	0,9	1,12	1,24	1,41	1,45	1,49	1,51	1,48	1,56

$$CR = CI \ / \ RI \tag{3}$$

If *CR*<0.1, the pairwise comparison matrix is consistent. Otherwise, pairwise comparisons should be checked and revised, and the above calculations should be repeated.

Step 5: Analysis of the scores

The highest value alternative is chosen as the best alternative.

3.2. TOPSIS

TOPSIS was developed by Hwang and Yoon in 1981 and is a method commonly used in real life multi-criteria decision problems. This method allows decision-makers to compare and sort alternatives. TOPSIS ranks the alternatives based on the maximum distance from the negative ideal solution, and minimum distance to the positive ideal solution. After all, the method chooses the closest alternative to the ideal solution. The method consists of 6 steps [31].

Step 1: Creating the decision matrix

In the rows of the decision matrix, the alternatives are listed and the criteria which affects the decision-making process are given in the columns.

Step 2: Creating the standard decision matrix

A standard decision matrix is created with Eq. 4:

$$r_{ij} = aij \,/\, \sqrt{\sum_{k=1}^{m} a_{kj}^2} \tag{4}$$

Step 3: Creating the weighted standard decision matrix

The weighted standard decision matrix is obtained by multiplying each weight value calculated for the criteria by the value of the relevant criterion in the standard decision matrix.

Step 4: Creating the positive ideal (A) and negative ideal (A⁻) solutions*

According to the assumption that the criteria show a tendency to monotonous increasing and monotonous decreasing, the maximum and minimum ones of the values in the weighted standard decision matrix are determined for obtaining the positive and negative ideal solution sets.

Step 5. Calculating the separation measures

The distance of the criteria values of each decision point in the matrix to the ideal and negative ideal solution is calculated by using Eq. 5. and Eq. 6:

$$S_i^* = \sqrt{\sum_{j=1}^{n} \left(v_{ij} - v_j^*\right)^2} \tag{5}$$

$$S_i^- = \sqrt{\sum_{j=1}^{n} \left(v_{ij} - v_j^-\right)^2} \tag{6}$$

Step 6: Calculating the relative closeness to ideal solution

The relative closeness to the ideal solution C_i^* is calculated by using the separation measures according to the Eq. 7:

$$C_i^* = \frac{S_i^-}{S_i^- + S_i^*} \tag{7}$$

C_i* is in the range 0-1. If C_i*=1, the corresponding decision point is absolutely close to the ideal solution. On the other hand, C_i*=0 represents the absolute closeness of the decision point to the negative ideal solution [5].

3.3. Artificial neural network

ANN is a processor that has a natural tendency to put into practice the stored information based on experience. The ANN is similar to the human brain in two respects: the information is obtained by the network through a learning process and the so-called synaptic weight between neurons is used to store information [28]. The emergence and development of this method is as follows: McCulloch and Pitts [52] are taken the first steps for the ANN by setting up a simple neural network with a small electrical circuit. This network is emerged by imitating the computational ability of the human brain. In 1949, Hebb [30] described the basic theory of learning in his book, The Organization of Behavior. Rosenblatt [63] found the perception of perceptron in 1958 is an important development for ANN. In 1969, Minsky and Papert [54] proved that the perceptron sensors could not solve the

XOR problem and suggested that two-layer feed-forward networks could be used. Rumelhart et al. [64] developed the back-propagation algorithm for multi-layered neural networks in 1986. After this process, progress has also been made so far [29,89].

A simple neural network structure consists of inputs, weights, aggregation function, activation function and output element (Figure 2). The inputs consist of numerical values from the outside of the system or from neurons. The effect of input to the cell on the network refers to weights. Aggregation function performs a linear combination of the inputs of the neuron by calculating the net input to the cell. The numerical sum of these two vectors gives the net input and it is sent to the activation function [28]. The activation function is a non-linear function that shows the network's data structure [22]. There are many activation functions such as sigmoid, hyperbolic tangent, logarithmic sigmoid and purelin.

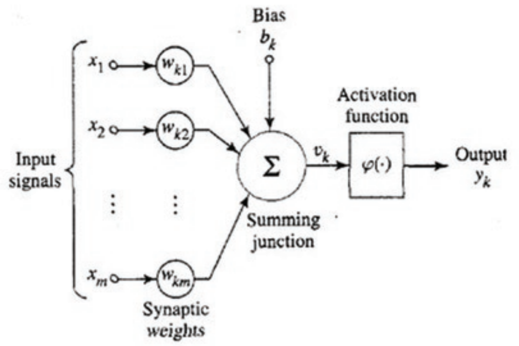


Fig. 2. A simple ANN structure [28]

ANN architecture is divided into single-layer and multi-layer artificial neural networks. The multi-layered structure is used in more complex problems and the single-layered structures are used in simpler problems. Due to the use of multi-layered networks in the complex problems of energy [74,83,85], multi-layered structure is utilized in this study.

4. Case Study

In particular, population growth, industrialization and urbanization with constantly evolving technology, the demand for electricity in Turkey is increased by an annual average rate of 5.6% in the last decade. At the same period, electricity consumption per capita in Turkey has reached from 2,052 kWh to 3,373 kWh with 64,4'% increase. Hydroelectric power plants met about one-fifth of electricity consumption, which realizes 292,171,618 kWh as of the end of 2018 [21], in Turkey with these significant increasing rates. Therefore, uninterrupted electricity generation in hydroelectric power plants has critical importance in terms of energy supply security for Turkey. Considering the fact that one of the two pillars of sustainable energy supply in electricity generation power plants is maintenance management, the critical importance of maintenance planning in hydroelectric power plants is seen.

In this context, especially the unwanted stoppages in the large-scale hydroelectric power plants and the damages caused by them have to be considered as important problems. Because, these stoppages affect the energy supply security of the country negatively, not only the plant owner organization. Considering that the power plant with 5 units of 200 MW and 6,111 equipment is operated in accordance with the operational directives, it is focused on the maintenance planning problem in this study, and it is reached the result of not carrying out the maintenance within a systematic plan is concluded with undesirable long-term stoppages in the plant.

It is not possible to apply a standard maintenance policy for all equipment in such a large plant, due to the high level of impact of maintenance on costs. For this reason, first, the most critical equipment groups which stop the electricity generation in the power plant and interfere with the generation quality and reliability are determined with AHP-TOPSIS combination. Based on the fact that the plant can provide sustainable energy supply with appropriate maintenance to be applied to these equipment groups, an ANN model is proposed by using 11-years fault data of them. As a result, possible fault periods of the most critical equipment of the power plant are obtained by using this model. Maintenance before the estimated breakdown dates is an accepted strategy that prevents the occurrence of faults, and consequently, a maintenance plan is prepared in line with this strategy and the results of the application are presented. The implementation stages of the study are presented in Figure 3.

4.1. Determining the critical equipment

At this stage, which is carried out according to the TOPSIS methodology, the evaluation criteria specified in Table 3 are determined first. The evaluation criteria are determined by referring to the opinions of the experts working in the plant and by taking into account all factors affecting the criticality level of each equipment for the plant and are determined to be related to each equipment.

In Table 3, the verbal value is assigned to each equipment according to each criterion using the parameters column specified under the criteria. The numerical equivalents of the parameters are created by utilizing the views of the power plant experts when considering the assumption of "all indicators must be numerical" for implementing the TOPSIS method. While determining the numerical equivalents of the parameters, a scale consisting of integers between 0-10 is used, and the highest score (10) is given to the parameter which directly affects the electricity generation in the power plant (unit shutdown). Other parameter scores are determined by considering the highest scores given between all the criteria and the scores given between each criterion.

Upon completion of this stage, the initial decision matrix which dimension is 6,111 x 9 is obtained and TOPSIS methodology is started. In order to determine the criticality levels of 6,111 equipment under 9

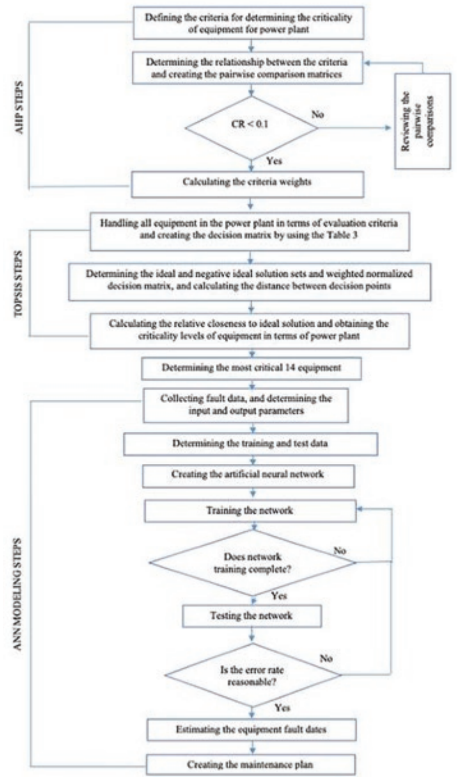


Fig. 3. Application steps

Table 3. Evaluation criteria

	Criteria	Criteria Parameters	Numerical Equivalents of the Parameters
C1	Warehouse backup	Never	3
		Sometimes	2
		All the time	1
C2	Maintenance pre-conditions	Unit shutdown	7
		Shutdown by situation	6
		Shutdown by time	5
		Maintenance without back up	2
		Shutdown does not require	1
C3	Additional work requirement	Required	5
		Not required	1
C4	Failure period	Monthly	8
		Quarterly	5
		Semi - annually	3
		Annually	2
		Long term	1
		Unknown	1
C5	Possible consequences	Unit shutdown	10
		Problem in emergency situation	9
		Load reduction	8
		Running without back up	7
		Equipment shutdown	6
		Security problem	6
		Deficient function	2
		Damage in associated equipment	2
		Problem in start	1
		Fluid consumption increase	1
C6	Availability of measuring equipment	Yes	3
		No	1
C7	Static, dynamic or electrical property of equipment	Mechanical-dynamic	2
		Mechanical-static	1
		Electrical	1
		I&C	1
C8	Fault shooting time	One week	9
		More than one day	3
		Unknown	3
		2-8 hours	2
		Less than 2 hours	1
C9	Detectability of failure	Difficult	3
		Easy	1

criteria in terms of power plant, the evaluation criteria weights are calculated with AHP at first. The CR of the pairwise comparison matrix formed between the criteria is calculated as 0.051, and it means the relevant matrix is consistent. The weight of the 9 criteria as a result of the calculations made on this consistent matrix are given in Table 4.

Weighted normalized decision matrix is formed by using the criteria weights calculated with AHP and, ideal and negative ideal solution sets are determined. Then, the separation measures are calculated for each equipment with Eq.5 and Eq.6. Finally, equipment priority levels (C_i^*), which are defined as relative closeness to ideal solution, of each equipment are found by using Eq.7.

According to the priority levels of the equipment calculated with TOPSIS, the most critical equipment in terms of power plant is found as turbines, generators and disconnectors with a value of 0.837 C_i^*.

Table 4. Criteria weights

Criteria		Weights
C1	Warehouse backup	0.051
C2	Maintenance pre-conditions	0.241
C3	Additional work requirement	0.029
C4	Failure period	0.071
C5	Possible consequences	0.400
C6	Availability of measuring equipment	0.062
C7	Static, dynamic or electrical property of equipment	0.055
C8	Fault shooting time	0.029
C9	Detectability of failure	0.062

This value is considered as 100 full score to ensure the ease of calculations and the scores of all remaining equipment are recalculated accordingly. As a result of this process, power plants experts have determined that the equipment which directly affects the sustainable power generation have 95 and more scores. In this context, the most critical 14 equipment in the power plant which determines the scope of the study are given in Table 5.

Table 5. The most critical equipment of the power plant

Rank	Equipment	Score
1	Turbine	100
2	Generator	100
3	Disconnector	100
4	Intake structure	98
5	Butterfly valve	98
6	Main power transformer	97.5
7	Brake system	97.5
8	Compressed oil tank	96.9
9	Cooling water structure	96.9
10	Wicket gate	96.8
11	Relay	96.7
12	Excitation transformer	96.7
13	Speed governor	95.9
14	Circuit breaker	95

4.2. Failure date estimation of the critical equipment

In this study, it is focused on the MTBF which is one of the maintenance performance parameters for determining the possible breakdown dates of the 14 most critical equipment of the power plant. On the first step of the prediction stage, the input parameters affecting the breakdown are determined by 8 power plant professionals, each having 10 to 25 years of hydroelectric power plant operation and maintenance experience and their occupations are industrial, electrical, electrical-electronic and mechanical engineer. Equipment type, pressure effect, economic life of the equipment, fault repair time and predictive maintenance effect are determined as input parameters. The output parameter is determined as the number of days between two faults of each equipment based on the fault data recorded since 2005. Since the numerical use of data in the ANN models has a positive effect on the education of the network, the data obtained are converted to numerical form. The data set used in the study is given in Annex-1. The process of distinguishing data is one of the factors affecting the education of the network. Because, with the application of different training and test data groupings, it has been observed that the test results have changed although the structure of the network is not changed. The success of artificial neural network application is closely related to the approaches and experiences to be applied. Determining the appropriate structure in the success of the application is another factor that greatly influences the results of the model. In this context, 80% (228 faults) of the 285 faults that occurred in 11 years were used for training in the network and the remaining part (57 faults) was used to test the performance of the network. The 228-training data are allocated with rates of 70-15-15 in MATLAB as training, validation and test data respectively. Weights are estimated during the training phase. The generalization ability of untrained data is preserved during the validation phase. In the testing phase, the error rate is calculated.

Multiple attempts should be made to find the appropriate results in the network model. These trials are carried out in three stages. First, the process is continued until the learning is achieved. In other words, trials are maintained until the deviations from the target values fall below a certain rate. If the proximity to the target values is provided, the learning is stopped for the network, and the samples which are not shown in learning phase are submitted to the network, and thus, test phase is started. If the deviation between the test results and the targeted values is not acceptable, improvements are made by backing to the learning stage of the network, and this process is continued until it is close to both learning and test objectives. Therefore, the process after the completion of the data related to the study is continued by examining the ANN models. The studies on ANN models are conducted with MATLAB program and the network model and algorithm to be used are investigated [55]. The network type, number of layers in the network, number of neurons in each layer, types of learning and activation function types used in network training, learning and momentum coefficients and the number of iterations are changed separately according to the results obtained during the study. The ANN architecture proposed in this study selected by comparison of network trial results as in most studies in the literature [11,14,28,46,60,71, etc.]. The suitability of the ANN model with 285 data was examined by comparing the results of approximately 150 trials rather than the Vapnik – Chervonenkis dimension theory. Learning and generalization errors of each ANN model were observed, and the model with the lowest error rate, and no memorization in the performance graph was chosen. As a result, it was concluded that 5-20-10-1 ANN model was suitable for solving this problem with 285 fault data that occurred in the 11-year time period in the hydroelectric power plant. The network structure is given in Figure 4.

Figure 5 shows the network with the best regression graph results among the networks established by changing the above-mentioned parameters. 2 hidden layers are used in this selected network, and the

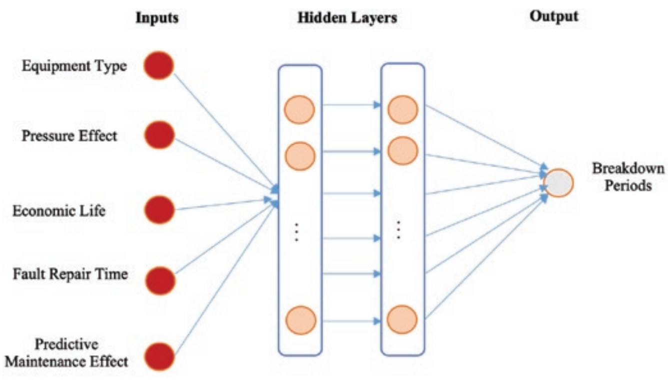


Fig. 4. Network structure

number of neurons in these layers are 20 and 10, respectively. In the network structure, purelin and tansig functions were used as transfer functions respectively, and learngdm was used as a learning function. Moreover, Levenberg-Marquardt algorithm was used in the training of the network. "Epoch=1000", "performance goal=0", "learning rate=0.01", "momentum constant=0.9" and "maximum validation failures=6" were taken as stopping criteria. These are the stopping criteria in the ANN module in the MATLAB program.

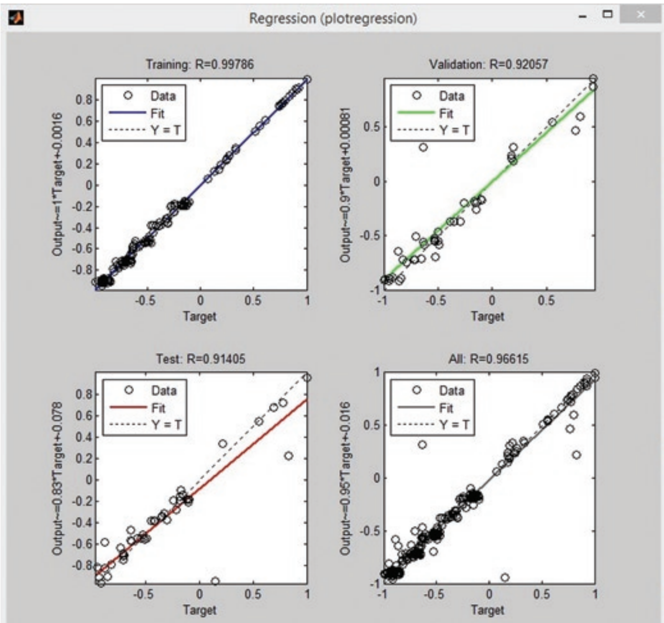


Fig. 5. Regression graphs of selected network

Table 6. Average fault periods of critical equipment

Rank	Equipment	Average Fault Period (Day)
1	Turbine	1,096
2	Generator	1,095
3	Disconnector	733
4	Intake structure	730
5	Butterfly valve	201
6	Main power transformer	1,091
7	Brake system	731
8	Compressed oil tank	367
9	Cooling water structure	365
10	Wicket gate	1,457
11	Relay	1,461
12	Excitation transformer	722
13	Speed governor	372
14	Circuit breaker	728

The performance of the network is tested with unused data in the training of the network. The Mean Absolute Percentage Error (MAPE) and R^2 values of the performance tested network are obtained as 0.03 and 0.91, respectively. The values are consistent with the performance values of the studies examined in the literature. Therefore, this network is selected for the failure date estimation problem handled in this study. As a result of the estimation phase, the breakdown periods are determined for each critical equipment. The average fault period for the equipment is given in Table 6. Finally, maintenance plans are applied in such a way that they will be performed before to the failure dates.

5. Conclusions and Recommendations

Electricity generation power plants are continuous production facilities focused on sustainable energy supply, which consist of continuity, reliability, efficiency, economy and environmental sensitivity. In order to achieve this comprehensive objective, it is necessary to comply with the operating rules determined by the power plant manufacturers and to implement the effectively managed maintenance processes simultaneously. In this context, it is a necessity to manage the maintenance processes to be carried out in the power plants, and the most critical and complex phase of maintenance management in electricity generation power plants, which consists of thousands of equipment, is maintenance planning. From this point of view in this study, a maintenance planning is performed with the integration of AHP-TOPSIS-ANN for 14 equipment with the highest importance in one of the large-scale hydroelectric power plants which meet about quintile of Turkey's energy demand.

This study has the feature of being the first in the literature in terms of both the use of these methods in an integrated manner and creating a maintenance schedule with MTBF estimation in hydroelectric power plants. In the studies examined in the literature, condition-based fault estimations are made for a single equipment or machine system. However, in this study, a preliminary qualification is performed for all equipment in a large-scale hydroelectric power plant and the most critical equipment are selected by analytical methods in wide perspective. In hydroelectric power plants where the maintenance activities are costly in terms of material, manpower and time requirements, and generation losses, these selected equipments are directly affecting the sustainable energy generation in the power plant. In this context, only the effective maintenance planning to be applied to these equipments meet the main target of the power plant and this corresponds exactly with the real-life plant management. The selected equipment is determined not only for a single service of the plant but also for its mechanical, electrical and electronic equipment. In this context, a structure covering the entire power plant is established. In addition to this, considering the lack of measuring sensors in all of the equipment, the possibility of arriving at very close time intervals of the signals and the difficulty of interfering with the equipment, and the costs about equipping all of the equipment with measuring sensors, the occurrence of failure etc., interfering with the equipment before the failure occurrence by using the signals of the sensors called as predictive maintenance have no effective results at the power plant. Because, prolonged and unexpected faults have often occurred, and the expected output from the plant could not be obtained for the plant owner and country. However, by performing the proposed maintenance planning approach within the scope of the study, significant improvements are achieved in the power plant in terms of increasing generation efficiency, extending the economic life of the power plant, minimization of generation and maintenance costs, maximization of availability ratio and profit maximization.

Hajian and Styles stated that in this state the output trains the targets very well for training, testing and validation and the R-value is over 0.9 also in this case the test set error and the validation set error have similar characteristics, so the ANN response is satisfactory [27]. Our ANN model is also suitable for explanation in this sentence. Furthermore, in the proposed model, two-stage test application and the error rate were found to be 0.03, which is an accepted rate in the literature. In addition, the purpose of the model is to ensure energy supply security by keeping downtime caused by malfunction to a minimum. When maintenance planning is made according to the model results, it has been observed that the stoppages caused by malfunctions are prevented and the results are effective in this sense.

In order to validate the maintenance schedule obtained as a result of this study and to determine the added-value, relevant maintenance schedule has been implemented in 1 unit of the power plant for 2 years. In this context, no unit shutdowns have occurred due to lack of maintenance in 14 selected equipment as a result of maintenance activities carried out the power plant. When the malfunction data are analyzed, it is determined that the time required to eliminate such a malfunction varies between 8 hours and 20 days. In other words, a minimum of an 8-hours shutdown has been prevented and a potential loss of 1.6 million kWh of energy has not occurred.

One of the maintenance strategies required for hydroelectric power plants is revision maintenance. This maintenance strategy needs to be implemented at an average of 2 years at the plant where the application is carried out for each unit, and an average of 110 calendar days is needed. The proposed maintenance schedule is also positively affected the revision maintenance performed. This is because only the necessary maintenance processes are carried out to include all stages of implementation in the power plant which has previously applied only the corrective maintenance strategy. Thus, the revision period is reduced to 82 days. This means a 34% improvement. Considering that the unit could not be used in electricity generation during the revision period, a total of 134.4 million kWh energy is gained for the 28-days difference between the pre- and post- period of proposed approach. These results show that the power plant owner gains millions of liras, as well as Turkey has an important contribution to the energy supply security.

As a result of the implementation of the entire maintenance schedule, the added-value will be increased as a result of the implementation of the said plan in the other units of the plant. The above-mentioned additives of the study can be considered as important contributions to the literature when evaluated with the bullets mentioned at the end of the second section.

In addition, as a continuation of this study, the entire process of the operation should be reevaluated over the fault-free operation period of the power plant and a new network design can be made that learns the effect of maintenance on the system as well as a period that learns the faults.

Acknowledgement

This study was conducted at Kırıkkale University Projects Coordination Unit coded 2018/008 scientific research project.

Annex 1

Table 7. Data set

Equipment	Pressure effect	Economic life of the equipment	Fault repair time	Predictive maintenance effect	Breakdown periods
Intake structure	81	1.550	288	0	960
Cooling water structure	56	3.300	281	4	937
Butterfly valve	80	1.753	257	12	856
Relay	6	4.002	238	52	298
Disconnector	3	2.675	232	0	772
Butterfly valve	80	1.341	230	0	287
Intake structure	72	1.723	228	1	761
Relay	4	3.010	226	0	282
Cooling water structure	63	3.006	214	4	268
Speed governor	81	2.941	212	4	708
Disconnector	1	0.677	208	12	694
Cooling water structure	70	2.151	203	0	254
Cooling water structure	70	2.013	198	0	248
Butterfly valve	90	1.505	198	0	659
Excitation transformer	12	1.317	191	0	239
Cooling water structure	70	3.437	190	4	237
Turbine	80	2.871	188	4	235
Generator	90	2.687	181	0	603
Speed governor	90	2.676	179	4	596
Intake structure	81	1.293	173	0	216
Excitation transformer	16	1.342	167	0	418
Speed governor	90	3.077	165	4	413
Disconnector	2	0.345	165	0	206
Wicket gate	12	4.018	164	12	2466
Generator	100	2.848	160	0	200
Disconnector	2	0.183	159	0	529

Relay	6	2.675	157	0	392
Turbine	80	2.443	156	0	520
Relay	10	2.848	153	0	383
Cooling water structure	56	2.430	152	0	506
Relay	6	2.337	147	0	35
Turbine	72	1.889	146	4	486
Compressed oil tank	72	1.670	145	2	2169
Main power transformer	12	2.866	143	52	2143
Butterfly valve	100	1.259	141	0	470
Speed governor	81	2.009	141	0	197
Circuit breaker	35	2.838	140	0	35
Turbine	72	2.159	133	0	333
Wicket gate	18	2.852	133	0	186
Butterfly valve	90	1.924	132	12	441
Wicket gate	12	2.842	131	0	184
Butterfly valve	90	1.838	131	12	437
Brake system	4	2.306	130	2	1957
Turbine	80	1.205	130	0	182
Turbine	80	1.550	129	0	322
Excitation transformer	20	1.424	128	0	320
Excitation transformer	12	1.508	127	0	318
Speed governor	81	3.204	127	4	178
Turbine	64	2.730	124	4	310
Speed governor	81	1.873	124	0	99
Butterfly valve	100	1.510	123	0	172
Turbine	64	1.978	121	4	303
Excitation transformer	20	1.923	120	4	1801
Circuit breaker	30	2.850	120	0	96
Relay	8	4.010	119	52	167
Butterfly valve	100	1.925	119	12	29
Generator	100	2.838	119	0	95
Speed governor	90	2.001	118	0	71
Relay	10	2.507	117	0	164
Speed governor	81	2.405	116	0	93
Generator	90	3.677	115	4	1730
Relay	10	3.838	115	52	1725
Wicket gate	6	2.345	114	0	91
Speed governor	72	3.079	113	4	29
Speed governor	72	2.401	112	0	157
Relay	8	3.950	111	52	1661
Main power transformer	12	2.009	110	0	154
Generator	90	2.507	109	0	87
Butterfly valve	100	1.171	108	0	151
Disconnector	1	2.338	108	0	86
Cooling water structure	63	3.444	107	4	150
Circuit breaker	25	3.010	105	0	123
Wicket gate	12	2.513	105	0	84
Wicket gate	18	2.682	104	0	146
Turbine	80	1.891	104	4	83

Main power transformer	6	2.870	103	52	82
Speed governor	81	2.144	100	0	140
Speed governor	90	2.277	96	0	77
Speed governor	90	1.868	96	0	26
Wicket gate	18	2.508	96	0	134
Cooling water structure	56	2.014	93	0	30
Excitation transformer	16	1.927	93	4	130
Circuit breaker	35	3.015	93	0	74
Relay	8	2.345	93	0	25
Wicket gate	12	2.673	91	0	128
Speed governor	81	2.271	90	0	126
Disconnector	1	0.678	90	12	29
Main power transformer	12	2.873	89	52	71
Compressed oil tank	80	1.171	88	0	123
Speed governor	81	2.273	88	0	70
Speed governor	90	2.412	87	0	31
Main power transformer	6	3.004	86	52	121
Intake structure	90	1.295	84	0	117
Turbine	64	3.157	84	4	117
Circuit breaker	35	2.505	83	0	99
Cooling water structure	63	2.866	82	4	1224
Speed governor	72	2.145	80	0	32
Turbine	64	2.293	79	0	111
Relay	10	2.352	79	0	110
Circuit breaker	25	2.517	78	0	29
Wicket gate	12	3.003	77	0	108
Main power transformer	6	3.436	77	52	1152
Speed governor	72	2.136	76	0	107
Circuit breaker	20	3.018	71	0	27
Wicket gate	6	2.518	70	0	84
Relay	6	2.680	70	0	24
Circuit breaker	30	2.675	68	0	82
Circuit breaker	25	3.020	65	0	15
Circuit breaker	25	2.520	60	0	29
Relay	6	2.682	60	0	31
Wicket gate	12	2.345	56	0	15
Circuit breaker	20	3.015	55	0	13
Intake structure	72	1.551	51	0	6
Intake structure	81	1.372	50	0	6
Intake structure	90	1.372	50	0	5
Cooling water structure	56	3.301	49	4	8
Circuit breaker	25	2.838	49	0	6
Relay	8	2.352	49	0	5
Circuit breaker	30	3.003	48	0	7
Circuit breaker	30	2.843	47	0	5
Circuit breaker	25	2.680	47	0	6
Circuit breaker	25	3.018	47	0	8
Circuit breaker	30	3.017	47	0	15
Circuit breaker	30	2.520	47	0	7

Circuit breaker	35	2.682	47	0	35
Circuit breaker	30	2.680	47	0	5
Circuit breaker	35	2.342	46	0	6
Circuit breaker	30	2.838	46	0	4
Circuit breaker	30	3.003	46	0	8
Circuit breaker	20	2.505	46	0	5
Circuit breaker	20	2.517	45	0	7
Circuit breaker	25	2.835	45	0	4
Wicket gate	18	2.520	45	0	21
Circuit breaker	25	2.340	45	0	6
Butterfly valve	80	1.924	45	12	5
Circuit breaker	25	2.683	44	0	4
Wicket gate	12	2.350	44	0	7
Speed governor	81	2.409	44	0	21
Circuit breaker	20	2.512	44	0	8
Butterfly valve	80	1.173	44	0	6
Circuit breaker	20	2.670	44	0	4
Speed governor	81	2.011	44	0	5
Speed governor	90	2.147	43	0	23
Speed governor	90	1.879	43	0	15
Circuit breaker	25	2.668	43	0	4
Speed governor	72	2.411	43	0	6
Speed governor	72	2.409	42	0	8
Circuit breaker	25	2.342	42	0	4
Speed governor	72	2.011	41	0	4
Circuit breaker	35	2.687	41	0	27
Speed governor	72	1.880	41	0	8
Intake structure	90	1.551	41	0	9
Speed governor	90	1.868	40	0	4
Circuit breaker	20	3.003	40	0	3
Circuit breaker	20	2.835	39	0	3
Speed governor	72	1.876	39	0	8
Circuit breaker	25	2.687	39	0	3
Circuit breaker	30	3.020	39	0	9
Circuit breaker	35	2.683	38	0	3
Circuit breaker	30	2.835	38	0	24
Circuit breaker	20	2.670	38	0	15
Cooling water structure	70	2.151	38	0	1
Main power transformer	9	3.436	38	52	8
Circuit breaker	35	2.342	38	0	3
Circuit breaker	20	2.352	37	0	27
Speed governor	81	2.409	37	0	3
Circuit breaker	30	2.845	37	0	9
Speed governor	81	2.011	36	0	3
Circuit breaker	25	2.852	36	0	18
Speed governor	72	2.009	36	0	3
Circuit breaker	30	2.683	35	0	9
Speed governor	90	1.876	35	0	3
Circuit breaker	20	3.018	35	0	1

Circuit breaker	20	2.843	35	0	28
Relay	6	2.345	35	0	10
Intake structure	72	1.296	34	0	14
Wicket gate	6	2.502	33	0	9
Circuit breaker	25	3.017	33	0	10
Circuit breaker	30	2.672	33	0	28
Circuit breaker	25	2.853	32	0	12
Circuit breaker	35	2.852	32	0	2
Speed governor	90	2.411	32	0	20
Circuit breaker	25	3.010	32	0	1
Circuit breaker	25	2.842	32	0	15
Circuit breaker	35	2.685	32	0	2
Circuit breaker	30	2.668	31	0	2
Circuit breaker	35	2.345	31	0	13
Circuit breaker	30	3.012	31	0	10
Circuit breaker	35	2.520	31	0	2
Circuit breaker	30	2.505	31	0	2
Circuit breaker	25	2.505	30	0	2
Speed governor	90	2.012	30	0	24
Circuit breaker	25	2.352	30	0	2
Speed governor	90	2.405	30	0	11
Circuit breaker	35	2.350	29	0	2
Speed governor	72	2.941	29	4	2
Circuit breaker	35	2.520	29	0	1
Speed governor	90	2.411	29	0	2
Speed governor	72	2.325	20	0	2
Circuit breaker	20	2.342	28	0	23
Speed governor	81	2.144	28	0	2
Speed governor	63	2.011	28	0	12
Speed governor	72	2.144	27	0	2
Brake system	3	1.802	27	0	2
Relay	8	3.950	27	52	19
Intake structure	90	1.296	27	0	17
Circuit breaker	20	2.853	26	0	2
Circuit breaker	25	2.345	26	0	1
Speed governor	72	2.409	25	0	1
Circuit breaker	20	2.348	25	0	28
Circuit breaker	25	2.342	25	0	18
Cooling water structure	63	2.013	25	0	13
Speed governor	90	2.277	23	0	1
Circuit breaker	30	2.343	23	0	1
Circuit breaker	20	2.345	23	0	10
Generator	100	2.502	21	0	19
Speed governor	72	2.273	21	0	1
Relay	4	2.682	21	0	0
Circuit breaker	30	3.003	20	0	60
Circuit breaker	20	2.668	20	0	28
Compressed oil tank	80	1.672	19	2	58
Relay	8	2.340	19	0	58

Intake structure	81	1.372	19	0	1
Speed governor	72	2.011	19	0	1
Circuit breaker	35	3.015	19	0	0
Circuit breaker	35	2.340	19	0	56
Cooling water structure	56	2.156	19	0	15
Wicket gate	12	2.337	18	0	54
Excitation transformer	12	1.923	18	4	1
Circuit breaker	20	2.508	18	0	54
Circuit breaker	30	2.337	18	0	54
Relay	4	4.010	18	52	0
Wicket gate	18	2.352	18	0	53
Cooling water structure	63	2.154	18	0	53
Turbine	80	1.717	18	4	53
Circuit breaker	35	2.683	17	0	14
Relay	8	2.678	20	0	52
Speed governor	63	2.011	15	0	1
Circuit breaker	30	2.838	17	0	51
Circuit breaker	35	2.510	17	0	51
Excitation transformer	16	1.425	17	0	50
Circuit breaker	35	2.347	16	0	13
Circuit breaker	35	2.680	16	0	21
Intake structure	72	1.372	16	0	0
Generator	90	3.677	15	4	1
Speed governor	90	2.148	15	0	20
Speed governor	81	1.880	15	0	45
Speed governor	90	1.869	15	0	22
Circuit breaker	30	2.685	15	0	23
Circuit breaker	20	2.340	15	0	0
Circuit breaker	25	2.350	15	0	12
Speed governor	90	2.013	15	0	44
Turbine	72	2.874	15	4	44
Turbine	72	1.372	15	0	44
Wicket gate	18	2.340	14	0	43
Circuit breaker	30	2.350	14	0	41
Circuit breaker	20	2.345	14	0	11
Wicket gate	6	2.502	13	0	40
Brake system	4	1.802	13	0	40
Speed governor	81	2.411	13	0	1
Cooling water structure	70	2.867	13	4	38
Speed governor	90	2.144	12	0	10
Speed governor	90	2.137	12	0	37
Circuit breaker	25	2.845	12	0	37
Speed governor	72	1.876	12	0	36
Circuit breaker	30	2.668	12	0	0
Circuit breaker	20	2.518	12	0	21
Speed governor	63	1.877	12	0	35
Circuit breaker	30	2.352	11	0	0
Circuit breaker	35	2.518	9	0	0
Circuit breaker	30	2.342	9	0	0

Circuit breaker	25	2.668	7	0	0
Relay	10	2.343	7	0	70
Disconnector	3	0.187	7	0	68
Circuit breaker	25	2.515	7	0	68
Wicket gate	18	2.348	7	0	66
Turbine	64	1.715	6	4	64
Wicket gate	6	2.685	6	0	62
Speed governor	72	2.415	6	0	62
Main power transformer	9	2.011	6	0	61
Speed governor	81	2.408	6	0	61
Butterfly valve	90	1.173	6	0	61
Circuit breaker	30	2.342	6	0	0
Speed governor	81	1.876	5	0	0

References

1. Adoghe AU, Awosope COA, Ekeh JC. Asset maintenance planning in electric power distribution network using statistical analysis of outage data. International Journal of Electrical Power & Energy Systems 2013; 47: 424-35, https://doi.org/10.1016/j.ijepes.2012.10.061.
2. Adya M, Collopy F. How effective are neural networks at forecasting and prediction? A review and evaluation. Journal of Forecasting 1998; 17: 481-95, https://doi.org/10.1002/(SICI)1099-131X(1998090)17:5/6<481::AID-FOR709>3.0.CO;2-Q.
3. Al-Shayea QK. Artificial neural networks in medical diagnosis. International Journal of Computer Science Issues 2011; 8: 150-154.
4. Aly MF, El-hameed HMA. Integrating AHP and genetic algorithm model adopted for personal selection. International Journal of Engineering Trends and Technology (IJETT) 2013; 6.
5. Arıbaş M, Özcan U. Akademik araştırma projelerinin AHP ve TOPSIS yöntemleri kullanılarak değerlendirilmesi. Politeknik Dergisi 2016; 19: 163-173.
6. Ayodeji A, Liu Y kuo, Xia H. Knowledge base operator support system for nuclear power plant fault diagnosis. Progress in Nuclear Energy 2018; 105: 42-50, https://doi.org/10.1016/j.pnucene.2017.12.013.
7. Bangalore P, Patriksson M. Analysis of SCADA data for early fault detection, with application to the maintenance management of wind turbines. Renewable Energy 2018; 115: 521-532, https://doi.org/10.1016/j.renene.2017.08.073.
8. Barros CP, Wanke P. An analysis of African airlines efficiency with two-stage TOPSIS and neural networks. Journal of Air Transport Management 2015; 44-45: 90-102, https://doi.org/10.1016/j.jairtraman.2015.03.002.
9. Behzadian M, Kazemzadeh RB, Albadvi A, Aghdasi M. PROMETHEE: A comprehensive literature review on methodologies and applications. European Journal of Operational Research 2010; 200: 198-215, https://doi.org/10.1016/j.ejor.2009.01.021.
10. Behzadian M, Khanmohammadi Otaghsara S, Yazdani M, Ignatius J. A state-of the-art survey of TOPSIS applications. Expert Systems with Applications 2012; 39: 13051-13069, https://doi.org/10.1016/j.eswa.2012.05.056.
11. Bevilacqua M, Braglia M. Analytic hierarchy process applied to maintenance strategy selection. Reliability Engineering & System Safety 2000; 70: 71-83, https://doi.org/10.1016/S0951-8320(00)00047-8.
12. Bi R, Zhou C, Hepburn DM. Detection and classification of faults in pitch-regulated wind turbine generators using normal behaviour models based on performance curves. Renewable Energy 2017; 105: 674-688, https://doi.org/10.1016/j.renene.2016.12.075.
13. Box GEP, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons; 2015.
14. Braglia M, Carmignani G, Frosolini M, Zammori F. Data classification and MTBF prediction with a multivariate analysis approach. Reliability Engineering & System Safety 2012; 97: 27-35, https://doi.org/10.1016/j.ress.2011.09.010.
15. Chang P-L, Chen Y-C. A fuzzy multi-criteria decision making method for technology transfer strategy selection in biotechnology. Fuzzy Sets and Systems 1994; 63: 131-139, https://doi.org/10.1016/0165-0114(94)90344-1.
16. Chen B, Matthews PC, Tavner PJ. Wind turbine pitch faults prognosis using a-priori knowledge-based ANFIS. Expert Systems with Applications 2013; 40: 6863-6876, https://doi.org/10.1016/j.eswa.2013.06.018.
17. Chen D, Fang Z, Wu L, Zhu X. Study on the Prediction of Complex Equipment El MTBF DGMW (p/q)(1, 1) Model Based on the Small Sample. Journal of Grey System 2013; 25.
18. Cheng B, Titterington DM. Neural networks: A review from a statistical perspective. Statistical Science 1994: 2-30, https://doi.org/10.1214/ss/1177010638.
19. Çamkoru AM, Sayın VO. Bakım maliyeti yönetimi, Mühendis ve Makine 2012; 53(635): 16-20.
20. Dougherty M. A review of neural networks applied to transport. Transportation Research Part C: Emerging Technologies 1995; 3: 247-260, https://doi.org/10.1016/0968-090X(95)00009-8.
21. Energy statistics. UCTEA. The Chamber of Electrical Engineers (www.emo.org.tr/ekler/5aa1f79e3852d94_ek.pdf).
22. Fernandez C, Soria E, Martin JD, Serrano AJ. Neural networks for animal science applications: Two case studies. Expert Systems with Applications 2006; 31:444-450, https://doi.org/10.1016/j.eswa.2005.09.086.
23. Gardner MW, Dorling SR. Artificial neural networks (the multilayer perceptron)-a review of applications in the atmospheric sciences. Atmospheric Environment 1998; 32: 2627-2636, https://doi.org/10.1016/S1352-2310(97)00447-0.
24. Govindan K, Jepsen MB. ELECTRE: A comprehensive literature review on methodologies and applications. European Journal of Operational

Research 2016; 250: 1-29, https://doi.org/10.1016/j.ejor.2015.07.019.
25. Granger CWJ, Terasvirta T. Modelling non-linear economic relationships. OUP Cat 1993.
26. Gul M, Celik E, Aydin N, Taskin Gumus A, Guneri AF. A state of the art literature review of VIKOR and its fuzzy extensions on applications. Applied Soft Computing 2016; 46: 60-89, https://doi.org/10.1016/j.asoc.2016.04.040.
27. Hajian A, Styles P. Application of soft computing and intelligent methods in geophysics. Springer International Publishing 2018, https://doi.org/10.1007/978-3-319-66532-0.
28. Haykin S. Neural networks: a comprehensive foundation. Prentice Hall PTR; 1994.
29. Hebb DO. The Organizations of Behavior: a Neuropsychological Theory. Lawrence Erlbaum; 1963.
30. Hebbs DG. The organization of behavior. Wiely Sons, New York, NY, USA 1949.
31. Hwang C-L, Yoon K. Methods for multiple attribute decision making. Multiple Attribute Decision Making 1981: 58-191, https://doi.org/10.1007/978-3-642-48318-9_3.
32. Illias HA, Chai XR, Bakar AHA, Mokhlis H. Transformer incipient fault prediction using combined artificial neural network and various particle swarm optimisation techniques. PLoS One 2015; 10: 1-17, https://doi.org/10.1371/journal.pone.0129363.
33. Ishizaka A, Labib A. Review of the main developments in the analytic hierarchy process. Expert Systems with Applications 2011; 38: 14336-14345, https://doi.org/10.1016/j.eswa.2011.04.143.
34. Jiang R, Huang R, Huang C. Modeling the effect of environmental conditions on reliability of wind turbines. Journal of Shanghai Jiaotong University (Science) 2016; 21: 462-466, https://doi.org/10.1007/s12204-016-1747-7.
35. Jones B, Jenkinson I, Yang Z, Wang J. The use of Bayesian network modelling for maintenance planning in a manufacturing industry. Reliability Engineering & System Safety 2010; 95: 267-277, https://doi.org/10.1016/j.ress.2009.10.007.
36. Kahraman C, Ruan D, Doğan I. Fuzzy group decision-making for facility location selection. Information Sciences 2003; 157: 135-153, https://doi.org/10.1016/S0020-0255(03)00183-X.
37. Kalogirou SA. Artificial neural networks in renewable energy systems applications: a review. Renewable and Sustainable Energy Reviews 2001; 5: 373-401, https://doi.org/10.1016/S1364-0321(01)00006-5.
38. Komal S, Sharma SP. Fuzzy reliability analysis of repairable industrial systems using soft-computing based hybridized techniques. Applied Soft Computing 2014; 24: 264-276, https://doi.org/10.1016/j.asoc.2014.06.054.
39. Kubler S, Robert J, Derigent W, Voisin A, Le Traon Y. A state-of the-art survey & testbed of fuzzy AHP (FAHP) applications. Expert Systems with Applications 2016; 65: 398-422, https://doi.org/10.1016/j.eswa.2016.08.064.
40. Kumar A, Sah B, Singh AR, Deng Y, He X, Kumar P, et al. A review of multi criteria decision making (MCDM) towards sustainable renewable energy development. Renewable and Sustainable Energy Reviews 2017; 69: 596-609, https://doi.org/10.1016/j.rser.2016.11.191.
41. Kumar R, Singal SK. Penstock material selection in small hydropower plants using MADM methods. Renewable and Sustainable Energy Reviews 2015; 52: 240-255, https://doi.org/10.1016/j.rser.2015.07.018.
42. Kuo RJ, Chi SC, Kao SS. A decision support system for selecting convenience store location through integration of fuzzy AHP and artificial neural network. Computers in Industry 2002; 47: 199-214, https://doi.org/10.1016/S0166-3615(01)00147-6.
43. Kusiak A, Li W. The prediction and diagnosis of wind turbine faults. Renewable Energy 2011; 36:16-23, https://doi.org/10.1016/j.renene.2010.05.014.
44. Lee JW, Kim SH. Using analytic network process and goal programming for interdependent information system project selection. Computers & Operations Research 2000; 27: 367-382, https://doi.org/10.1016/S0305-0548(99)00057-X.
45. Liang R-H, Hsu Y-Y. Scheduling of hydroelectric generations using artificial neural networks. IEE Proceedings-Generation, Transmission and Distribution 1994;141: 452-458, https://doi.org/10.1049/ip-gtd:19941156.
46. Liao S-H, Wen C-H. Artificial neural networks classification and clustering of methodologies and applications-literature analysis from 1995 to 2005 Expert Systems with Applications 2007; 32: 1-11, https://doi.org/10.1016/j.eswa.2005.11.014.
47. Liberatore MJ, Nydick RL. The analytic hierarchy process in medical and health care decision making: A literature review. European Journal of Operational Research 2008; 189: 194-207, https://doi.org/10.1016/j.ejor.2007.05.001.
48. Liu Y, Fan J, Li Y. One system reliability assessment method for CNC grynder. Eksploatacja i Niezawodnosc - Maintenancen and Reliability 2014; 16(1): 97-104.
49. Lu Y, Sun L, Zhang X, Feng F, Kang J, Fu G. Condition based maintenance optimization for offshore wind turbine considering opportunities based on neural network approach. Applied Ocean Research 2018; 74: 69-79, https://doi.org/10.1016/j.apor.2018.02.016.
50. Mardani A, Zavadskas EK, Khalifah Z, Zakuan N, Jusoh A, Nor KM, et al. A review of multi-criteria decision-making applications to solve energy management problems: Two decades from 1995 to 2015. Renewable and Sustainable Energy Reviews 2017; 71: 216-256, https://doi.org/10.1016/j.rser.2016.12.053.
51. Mayadevi N, Vinodchandra S S, Ushakumari S. A review on expert system applications in power plants. International Journal of Electrical and Computer Engineering 2014; 4(1): 116.
52. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 1943; 5: 115-133, https://doi.org/10.1007/BF02478259.
53. Messai A, Mellit A, Abdellani I, Massi Pavan A. On-line fault detection of a fuel rod temperature measurement sensor in a nuclear reactor core using ANNs. Progress in Nuclear Energy 2015; 79: 8-21, https://doi.org/10.1016/j.pnucene.2014.10.013.
54. Minsky M, Paper S.Perceptrons. MIT Press, Cambridge MA, USA 1969.
55. Mohd Ali J, Hussain MA, Tade MO, Zhang J. Artificial Intelligence techniques applied as estimator in chemical process systems - A literature survey. Expert Systems with Applications 2015; 42: 5915-5931, https://doi.org/10.1016/j.eswa.2015.03.023.
56. Nasiri S, Khosravani MR, Weinberg K. Fracture mechanics and mechanical fault detection by artificial intelligence methods: A review. Engineering Failure Analysis 2017; 81: 270-293, https://doi.org/10.1016/j.engfailanal.2017.07.011.
57. Polo FAO, Ferrero Bermejo J, Gómez Fernández JF, Crespo Márquez A. Failure mode prediction and energy forecasting of PV plants to assist dynamic maintenance tasks by ANN based models. Renewable Energy 2015; 81: 227-238, https://doi.org/10.1016/j.renene.2015.03.023.
58. Özcan EC, Ünlüsoy S, Eren T. A combined goal programming - AHP approach supported with TOPSIS for maintenance strategy selection in hydroelectric power plants. Renewable and Sustainable Energy Reviews 2017; 78: 1410-1423, https://doi.org/10.1016/j.rser.2017.04.039.

59. Pankratz A. Forecasting with univariate Box-Jenkins models: Concepts and cases. vol. 224. John Wiley & Sons; 2009.
60. Ramadevi R, Sheela Rani B, Prakash V. Role of hidden neurons in an elman recurrent neural network in classification of cavitation signals. Int J Comput Appl 2012; 37(7): 9-13, https://doi.org/10.5120/4618-6626.
61. Ravì D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, Yang GZ. Deep learning for health informatics. IEEE journal of biomedical and health informatics 2016; 21(1): 4-21, https://doi.org/10.1109/JBHI.2016.2636665.
62. Ripley BD. Statistical aspects of neural networks. Networks and Chaos-Statistical and Probabilistic Aspects 1993; 50: 40-123, https://doi.org/10.1007/978-1-4899-3099-6_2.
63. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological Review 1958; 65: 386, https://doi.org/10.1037/h0042519.
64. Rumelhart DE, Hinton GE, Williams RJ. Learning internal representation by backpropagating errors. DE Rumelhart, & JL McCleland Parallel Distributed Processing: Explorations in the Microstructure of Cognition 1986; 1, https://doi.org/10.7551/mitpress/5236.001.0001.
65. Saaty TL. The analytic hierarchy process: planning, priority setting, resource allocation. McGraw-Hill International Book Company; 1980.
66. San Cristóbal JR. Multi-criteria decision-making in the selection of a renewable energy project in spain: The Vikor method. Renewable Energy 2011; 36: 498-502, https://doi.org/10.1016/j.renene.2010.07.031.
67. Schlechtingen M, Ferreira Santos I. Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection. Mechanical Systems and Signal Processing 2011; 25: 1849-1875, https://doi.org/10.1016/j.ymssp.2010.12.007.
68. Schlechtingen M, Santos IF. Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 2: Application examples. Applied Soft Computing 2014; 14: 447-460, https://doi.org/10.1016/j.asoc.2013.09.016.
69. Shankaracharya DO, Samanta S, Vidyarthi AS. Computational intelligence in early diabetes diagnosis: A review. The Review of Diabetic Studies: RDS 2010; 7: 252-261, https://doi.org/10.1900/RDS.2010.7.252.
70. Sharda R. Neural networks for the MS/OR analyst: An application bibliography. Interfaces (Providence) 1994; 24: 116-130, https://doi.org/10.1287/inte.24.2.116.
71. Sheela K G. Deepa S. N. Review on methods to fix number of hidden neurons in neural networks. Mathematical Problems in Engineering 2013, https://doi.org/10.1155/2013/425740
72. Sindhu S, Nehra V, Luthra S. Investigation of feasibility study of solar farms deployment using hybrid AHP-TOPSIS analysis: Case study of India. Renewable and Sustainable Energy Reviews 2017; 73: 496-511, https://doi.org/10.1016/j.rser.2017.01.135.
73. Sipahi S, Timor M. The analytic hierarchy process and analytic network process: an overview of applications. Management Decision 2010; 48: 775-808, https://doi.org/10.1108/00251741011043920.
74. Suganthi L, Samuel AA. Energy models for demand forecasting - A review. Renewable and Sustainable Energy Reviews 2012; 16:1223-1240, https://doi.org/10.1016/j.rser.2011.08.014.
75. Sun P, Li J, Wang C, Lei X. A generalized model for wind turbine anomaly identification based on SCADA data. Applied Energy 2016; 168: 550-567, https://doi.org/10.1016/j.apenergy.2016.01.133.
76. Taha Z, Rostam S. A fuzzy AHP-ANN-based decision support system for machine tool selection in a flexible manufacturing cell. The International Journal of Advanced Manufacturing Technology 2011; 57: 719-733, https://doi.org/10.1007/s00170-011-3323-5.
77. Vaidya OS, Kumar S. Analytic hierarchy process: An overview of applications. European Journal of Operational Research 2006; 169: 1-29, https://doi.org/10.1016/j.ejor.2004.04.028.
78. Vedachalam N, Ramadass GA. Reliability assessment of multi-megawatt capacity offshore dynamic positioning systems. Applied Ocean Research 2017; 63: 251-261, https://doi.org/10.1016/j.apor.2017.02.001.
79. Velasquez M, Hester PT. An analysis of multi-criteria decision making methods. International Journal of Operations Research 2013;10: 56-66.
80. Voyant C, Notton G, Kalogirou S, Nivet M-L, Paoli C, Motte F, et al. Machine learning methods for solar radiation forecasting: A review. Renewable Energy 2017; 105: 569-582, https://doi.org/10.1016/j.renene.2016.12.095.
81. Wang G, Huang SH, Dismukes JP. Product-driven supply chain selection using integrated multi-criteria decision-making methodology. International Journal of Production Economics 2004; 91: 1-15, https://doi.org/10.1016/S0925-5273(03)00221-4.
82. Wang JJ, Jing YY, Zhang CF, Zhao JH. Review on multi-criteria decision analysis aid in sustainable energy decision-making. Renewable and Sustainable Energy Reviews 2009; 13: 2263-2278, https://doi.org/10.1016/j.rser.2009.06.021.
83. Wang Z, Srinivasan RS. A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. Renewable and Sustainable Energy Reviews 2017; 75: 796-808, https://doi.org/10.1016/j.rser.2016.10.079.
84. Wanke P, Azad MAK, Barros CP, Hassan MK. Predicting efficiency in Islamic banks: An integrated multicriteria decision making (MCDM) approach. Journal of International Financial Markets, Institutions and Money 2016; 45: 126-141, https://doi.org/10.1016/j.intfin.2016.07.004.
85. Weron R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. International Journal of Forecasting 2014; 30: 1030-1081, https://doi.org/10.1016/j.ijforecast.2014.08.008.
86. White H. Learning in artificial neural networks: A statistical perspective. Neural Computation 1989; 1: 425-464, https://doi.org/10.1162/neco.1989.1.4.425.
87. Wong BK, Selvi Y. Neural network applications in finance: a review and analysis of literature (1990-1996). Information & Management 1998; 34:129-139, https://doi.org/10.1016/S0378-7206(98)00050-0.
88. Yang Y, Lu Z, Luo X, Ge Z, Qian Y. Mean failure mass and mean failure repair time: parameters linking reliability, maintainability and supportability. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014;16(2): 307-312.
89. Zhang G, Patuwo BE, Hu MY. Forecasting with artificial neural networks:: The state of the art. International Journal of Forecasting 1998; 14: 35-62, https://doi.org/10.1016/S0169-2070(97)00044-7.
90. Zhang GP. Neural networks for classification: a survey. IEEE Transactions on Systems, Man, and Cybernetics, Cybern Part C (Applications and Reviews) 2000; 30: 451-462, https://doi.org/10.1109/5326.897072.
91. Zhang Y, Ding X, Liu Y, Griffin PJ. An artificial neural network approach to transformer fault diagnosis. IEEE Transactions on Power Delivery 1996; 16: 55, https://doi.org/10.1109/MPER.1996.4311026.

92. Zyoud SH, Fuchs-Hanusch D. A bibliometric-based survey on AHP and TOPSIS techniques. Expert Systems with Applications 2017; 78:158-181, https://doi.org/10.1016/j.eswa.2017.02.016.

Evrencan ÖZCAN
Tuğba DANIŞAN
Rabia YUMUŞAK
Tamer EREN
Industrial Engineering Department
Kırıkkale University
71451, Kırıkkale, Turkey

E-mails: evrencan.ozcan@kku.edu.tr, tugbadanisan@gmail.com, rabiayumusak95@gmail.com, teren@kku.edu.tr

Article citation info:
SHAO Y, LIU B, WANG S, XIAO P. A novel test case prioritization method based on problems of numerical software code statement defect prediction. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 419–431, http://dx.doi.org/10.17531/ein.2020.3.4.

Yuanxun SHAO
Bin LIU
Shihai WANG
Peng XIAO

A NOVEL TEST CASE PRIORITIZATION METHOD BASED ON PROBLEMS OF NUMERICAL SOFTWARE CODE STATEMENT DEFECT PREDICTION

NOWATORSKA METODA PRIORYTETYZACJI PRZYPADKÓW TESTOWYCH OPARTA NA PROGNOZOWANIU BŁĘDÓW INSTRUKCJI KODU OPROGRAMOWANIA NUMERYCZNEGO

Test case prioritization (TCP) has been considerably utilized to arrange the implementation order of test cases, which contributes to improve the efficiency and resource allocation of software regression testing. Traditional coverage-based TCP techniques, such as statement-level, method/function-level and class-level, only leverages program code coverage to prioritize test cases without considering the probable distribution of defects. However, software defect data tends to be imbalanced following Pareto principle. Instinctively, the more vulnerable the code covered by the test case is, the higher the priority it is. Besides, statement-level coverage is a more fine-grained method than function-level coverage or class-level coverage, which can more accurately formulate test strategies. Therefore, we present a test case prioritization approach based on statement software defect prediction to tame the limitations of current coverage-based techniques in this paper. Statement metrics in the source code are extracted and data pre-processing is implemented to train the defect predictor. And then the defect detection rate of test cases is calculated by combining the prioritization strategy and prediction results. Finally, the prioritization performance is evaluated in terms of average percentage faults detected in four open source datasets. We comprehensively compare the performance of the proposed method under different prioritization strategies and predictors. The experimental results show it is a promising technique to improve the prevailing coverage-based TCP methods by incorporating statement-level defect-proneness. Moreover, it is also concluded that the performance of the additional strategy is better than that of max and total, and the choice of the defect predictor affects the efficiency of the strategy.

Keywords: *software defect prediction, test case prioritization, code statement metrics, machine learning, software testing.*

Metodę priorytetyzacji przypadków testowych (TCP) wykorzystuje się powszechnie do ustalania kolejności implementacji przypadków testowych, co przyczynia się do poprawy wydajności i alokacji zasobów w trakcie testowania regresyjnego oprogramowania. Tradycyjne techniki TCP oparte na pokryciu na poziomie instrukcji, metody/funkcji oraz klasy, wykorzystują pokrycie kodu programu tylko w celu ustalenia priorytetów przypadków testowych, bez uwzględnienia prawdopodobnego rozkładu błędów. Jednak dane o błędach oprogramowania są zwykle niezrównoważone zgodnie z zasadą Pareto. Instynktownie, im bardziej wrażliwy jest kod pokryty przypadkiem testowym, tym wyższy jest jego priorytet. Poza tym, pokrycie na poziomie instrukcji jest bardziej szczegółową metodą niż pokrycie na poziomie funkcji lub pokrycie na poziomie klasy, które mogą dokładniej formułować strategie testowe. Dlatego w artykule przedstawiamy podejście do priorytetyzacji przypadków testowych oparte na prognozowaniu błędów instrukcji oprogramowania, które pozwala zmniejszyć ograniczenia obecnych technik opartych na pokryciu. Wyodrębniono metryki instrukcji w kodzie źródłowym i zaimplementowano wstępne przetwarzanie danych w celu nauczania predyktora błędów. Następnie obliczono wskaźnik wykrywania błędów w przypadkach testowych poprzez połączenie strategii priorytetyzacji i wyników prognozowania. Wreszcie, oceniono wydajność ustalania priorytetów pod względem średnich procentowych błędów wykrytych w czterech zestawach danych typu open source. Kompleksowo porównano wydajność proponowanej metody w ramach różnych strategii ustalania priorytetów i predyktorów. Wyniki eksperymentów pokazują, że jest to obiecująca technika poprawy dominujących metod TCP opartych na pokryciu poprzez włączenie podatności na błędy na poziomie instrukcji. Ponadto stwierdzono również, że strategia dodatkowa cechuje się lepszą wydajnością niż strategie max i total, a wybór predyktora błędów wpływa na skuteczność strategii.

Słowa kluczowe: *przewidywanie błędów oprogramowania, priorytetyzacja przypadków testowych, metryki instrukcji kodu, uczenie maszynowe, testowanie oprogramowania.*

1. Introduction

Software testing, as an indispensable stage in software development life cycle, aims to discover defects in a software artefact as much as possible to assure its quality and reliability [10]. It is a very costly and resources consuming task that accounts for higher than 50% of development costs [15]. In previous work [51], test case selection (TCS), test case minimization (TCM) and TCP are three techniques maximize the value of test suites. TCP is often used to improve the effectiveness of some performance objectives associated with regression testing [7, 10, 21, 26, 35, 36], but it does not increase software costs like TCS and TSM by ignoring some meaningful test cases [9]. It prioritizes the test suite used according to certain test goals and strategies to optimize the speed of the defect detection rate, which is conducive to improving testing efficiency and reducing time and resource costs. With the expansion of application scenarios, it has been successfully applied in different test areas, such as graphical user interface testing [7, 20], web application testing [13].

TCP can schedule the execution of test cases in a specific order in order to enhance the defect detection rate, that is, to discover most defects as early as possible by prioritizing the most relevant test cases. It can be roughly grouped into multiple major dimensions [21, 51], such as requirement-based TCP [1, 22, 41], search-based TCP [26], coverage-based TCP [35], history-based TCP [40]. For instance, some researchers have studied the impact of requirement-based or search-based TCP methods on improving defect detection rates. For example, Arafeen et al. [1] and Krishnamoorthi et al. [22] showed that using requirement information in test case sequencing can improve testing efficiency. There are two common coverage strategies: total and additional. Li et al. [26] employed search-based heuristics TCP approaches (e.g.2-optimal greedy, hill climbing, genetic algorithms) to deal with the NP-hard problem for regression testing. Compared results on Siemens suite and space programs showed that genetic algorithms perform well. In addition, Spieker et al. [40] used neural networks and reinforcement learning to automatically select and prioritize test cases in continuous integration testing and tamed history-based TCP adaption problem. Among all TCP approaches, the coverage-based prioritization approach is one of the most commonly used in TCP [8, 17, 21, 27, 35, 36], such as statement-level coverage, branch-level coverage, method/function-level coverage and class-level coverage. The structural coverage-based method usually prioritizes test cases based on coverage, such as the total number of classes or methods covered. For example, Rothermel et al. [11, 35, 36] proposed a series of TCP methods including additional and total strategies combined with code coverage information. Henard et al. [17] comprehensively compared well-established white-box (e.g. total statement, additional branch) TCP techniques with black-box (e.g. input model diversity) ones. It is found that defects detected by the two techniques have high overlap and small performance differences. However, the black-box TCP techniques are more suitable for testing without source code. On the contrary, coverage objects considered in white-box TCP are source code, which is also the object concerned in this paper.

Unlike software reliability models [33, 38], software defect prediction can forecast defect prone or the number of defects in a software system, which is conducive to allocate testing resources efficiently. Furthermore, previous work [29, 32, 34, 43, 46, 48, 50] has suggested that using the defect-proneness of a defect predictor to rank makes sense for test cases. In other words, the more the code tends to finds defects, the higher the defect detection rate. Current defect predictors are mainly oriented to modules (e.g. class level [16], method/function level [25, 28, 37, 39, 44]), which are a relatively coarse granularity for TCP. For instance, Tonella et al. [43] integrated user knowledge through the case-based ranking machine learning algorithm with multiple, diverse TCP indexes. The method can handle partial, inconsistent and high-level data in a low-cost knowledge acquisition manner. The preliminary results showed that it is close to the optimal strategy for moderate test suite size (no more than 60). Wang et al. [7] presented quality-aware test case prioritization which leverages an unsupervised statistic defect prediction (CLAMI [30]) and a static bug finder (FindBugs [19]) to detect defective code and then revise the existing coverage-based methods by considering the weighted detective code. Empirical results performed on 7 open-source Java projects showed that it could improve coverage-based methods for test cases. Xiao et al. [48] designed a clustering-based TCP approach that utilized the defect-proneness probability based on the results of a method-level defect prediction model. The approach used a support vector machine to build the defect-proneness prediction model and employed K-means to calculate the optimal number of clusters. However, it only considered four code metrics: Line of code, total operators, total operands and cyclometric complexity. Lachmann et al. [23] proposed a system-level black-box TCP approach based on a support vector machine, which utilized test case history and natural language-based test case descriptions to prioritize. The method is compared with random and manual prioritization on the two subject systems, which show that it is beneficial to improve the defect detection rate. Mirarab et al. [29] provided a TCP approach based on the class-level Bayesian networks defect prediction model, which integrated software defect-proneness, code modification information, and test coverage data. The obtained results on a Java application showed that the approach has better test performance when there are a reasonable number of defects. Paterson et al. [32] proposed a TCP strategy (G-clef) based on defect prediction (Schwa) to reorder a test suite. They utilized code attributes such as the number of authors and revisions to configure Schwa, and compared G-clef on three groups of strategies: single-version, test execution history and software history. The results revealed that applying defect prediction to rank test cases was appealing.

However, as mentioned above, most of the current work focused on TCP methods based on coarse-grained software defect prediction models, such as class-level [29], method/function-level [7, 48], system-level [23]. Moreover, some techniques [32, 48] leverage only a small amount of metric information or user experience [43] to build defect-proneness models. If the granularity can be subdivided into the code statement-level, more accurate test resource allocation or case prioritization can be formulated, such as white-box TCP based on statement coverage [21, 35, 36]. However, these methods rank test cases in terms of the total number of statements covered without considering the probable distribution of code defects, that is, assuming that defects are evenly distributed in the source code. Nevertheless, the distribution of defects is skewed in most cases, with approximately 80% of defects occurring in 20% of the code [5]. In this paper, the defect-proneness probability of all valid statements in the code is predicted to guide coverage-based TCP. Therefore, integrating statement-based software defect prediction into a coverage-based test case prioritization has great theoretical and practical engineering value.

Furthermore, to the best of our present knowledge, there is no software defect prediction model that can predict defect prone of code statements. The key issue is how to select and extract statement metrics from source code. Besides, the existing mainstream research [29, 32, 34, 43, 46, 48, 50] only incorporates a single predictor into the TCP methods and does not compare the effects of various predictors on defect detection rate of test cases.

To overcome these shortcomings, we put forward a novel test case prioritization method based on software code statement defect prediction (TCP-SCSDP). The main idea of our method TCP-SCSDP is first to inject defects and extract code statement features to form a labelled defect dataset and perform data pre-processing on it. And then a supervised machine learning algorithm (e.g. random forest) is selected to build a software defect prediction model based on code statements. The model can assess the defect probability of statements and achieve the fine-grained prediction at the statement-level. Taking the white box test cases based on code coverage as the subject (test set), the prioritization strategy is utilized to calculate the defect detection rate of each test case. Finally, the weighted average percentage of defect detection is utilized to measure the ranking results. It can be observed that the code statement-level defect prediction method can not only improve the prediction granularity problem, but also fill the gap of the test case prioritization method based on the code statement-level defect-proneness prediction.

Our contributions to the current research are as follows:

(1) We propose a software defect prediction model based on code statement-level, which addresses the problems of acquiring statement metrics and prediction granularity.
(2) We incorporate code statement level software defect prediction into the test case prioritization process, which improves the defect detection rate.
(3) We systematically compare the effects of four common prioritization strategies and three classic predictors on test performance and find that the additional strategy is to the max and total ones and the choice of predictors has an impact on the prioritization strategy.
(4) We present an empirical evaluation using four open-source projects from the Software Infrastructure Repository (SIR). The experimental results indicate that our proposed TCP-SCS-DP is effective and feasible.

The remainder of this paper is organized as follows: The remainder of this paper is organized as follows: Section 2 presents the test case prioritization based on code statement defect prediction. Section 3 is devoted to the experimental setup. Experimental results and discussion are described in Section 4. Some threats to validity are described in Section 5. Finally, conclusions and future work are generalized in Section 6.

2. Test case prioritization based on code statement defect prediction

2.1. The proposed overall framework

The proposed method TCP-SCSDP is divided into two parts, code statement level defect prediction, and the prioritization strategy. The framework is presented in Fig. 1. The software defect prediction method based on code statement-level is a more fine-grained prediction than the traditional method based on function level or class level, which helps to solve the problem of unable to predict accurately. Firstly, metrics and defect information of code statements are extracted from the source code and software test report to form a labeled training dataset. And selecting a supervised learning algorithm (e.g. RF, GBRT, LRM) to train a software defect prediction model. Similarly, the features of the new software code statements are used as the validation test dataset. Secondly, the proposed prediction model is used to forecast the defect proneness of each line of code statement. Then we apply the test case priority such as total strategy and additional strategy [35] to calculate the defect probability of each test case. Finally, the test cases are sorted in terms of the probability of defects.

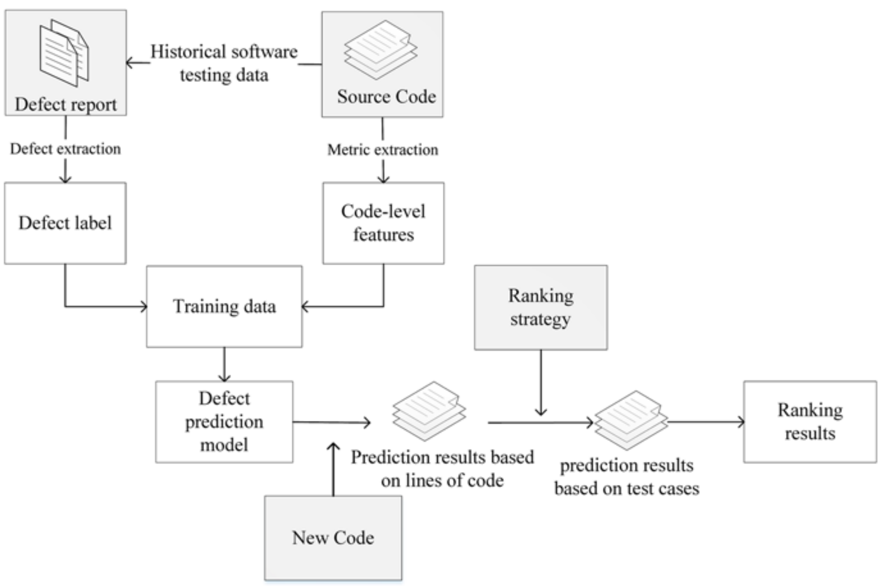


Fig. 1. The proposed test case prioritization framework based on statement-level defect prediction

2.2. Software defect prediction based on code statement level

2.2.1. Description of program code statement features

A key issue in statement-based software defect prediction is the way to extract code features. Traditional code feature (attribute or metric) such as McCabe, Halsted, and other metrics describes a module static metric information [28]. And machine learning algorithms are utilized to build the relationship between software module metrics and defect. The main difference from the traditional code features is that code vocabulary is used as independent variables in the software defect prediction based on code statement.

Regardless of the programming language, source code is composed of three vocabularies: language keywords, operators and operands.

- Language keywords, called reserved words, indicate an identifier that has a special meaning defined in advance by programming language. They can be used to describe a data type, the logical structure of a control program, and so on. In general, they cannot be used arbitrarily as variable names, method names, class names, package names, and parameters.

Table 1. 32 C language keywords defined by ANSI-C.

Type	Description
Type keywords	Basic type: void, char, int, short, long, double, float, signed, unsigned Complex type: enum, struct, union
Control keywords	Cycle control: for, do, while Condition control: if, else, switch, case, default Jump control: break, continue, goto, return
Storage keywords	auto, extern, register, static
Modifier keywords	const, sizeof, typedef, volatile

- Operators are symbols that represent specific operations and can be used to construct program language expressions. Common operators are mainly divided into five categories: arithmetic operators, relational operators, logical operators, assignment operators, as well as operators that complete bit operations - bitwise operators.
- An operand is an entity on which an operator acts and specifies the number operation variable in the instruction.

For example, the American National Standards Institute (ANSI) defines 32 language keywords for C programs. According to the characteristics of the programming language, four keywords are designed: type, control, data storage, and other keywords, among which type keywords indicate the type of data, control keywords are responsible for the logic of the processing program, store keywords declare variable range. Some of these types can be refined in combination with the functional struc-

Table 2. List of C language operators

Arithmetic operators	Relational operators	Logical operators	Bitwise operators	Assignment operators
+:addition or unary plus -:subtraction or unary minus *:multiplication /:division %:remainder ++:prefix increment –:prefix decrement	==:equal to !=:not equal to >:greater than >=:greater than or equal to <:less than <=:less than or equal to	&&: logical AND ||:logical OR !:logical NOT	&:bitwise AND |:bitwise OR :bit-wise complement ~:bitwise exclu-sive OR <<:left shift >>:right shift	=:basic assignment +=:addition assignment -=:subtraction assignment *=:multiplication assignment /=:division assignment %=:remainder assignment &=:bitwise AND assignment |=:bitwise OR assignment ~=:bitwise XOR assignment <<=:left shift assignment >>=: right shift assignment

ture. As shown in Table 1, control keywords can be divided into loop control, condition control, and jump control.

Operators are binary operators that assign the right operand to the left. Operators in C language are mainly divided into five categories, as shown in Table 2.

Obviously, there may be differences between keywords and operators for different programming languages, which are related to its own characteristics and design style. Java keywords are shown in Table 3. By comparing the keywords of C and Java language, we can see that C is a structured oriented language and Java is a typical object-oriented programming language. Therefore, there are many keywords describing classes in Java, such as abstract and interface, which are impossible to appear in C. In addition, although Java provides keywords such as try and catch for exception handling mechanisms, there is little difference between C and Java in a logic structure such as control and type keywords.

Table 3. List of Java language keywords

Keywords	Description
Access Modifiers	Private, protected, public
Class, Method, and	Abstract, class, extends, final, implements, interface, native, new
Variable Modifiers	Static, strictfp, synchronized, transient, volatile
Process control	Break, continue, return, do-while, if-else, for, instanceof, switch-case-default
Exception handling	Try catch, throw, throws
Package	Import, package
basic data types	Boolean, byte, char, double, float, int, long, short, null, true, false
variable reference	Super, this, void
reserved word	Goto, const

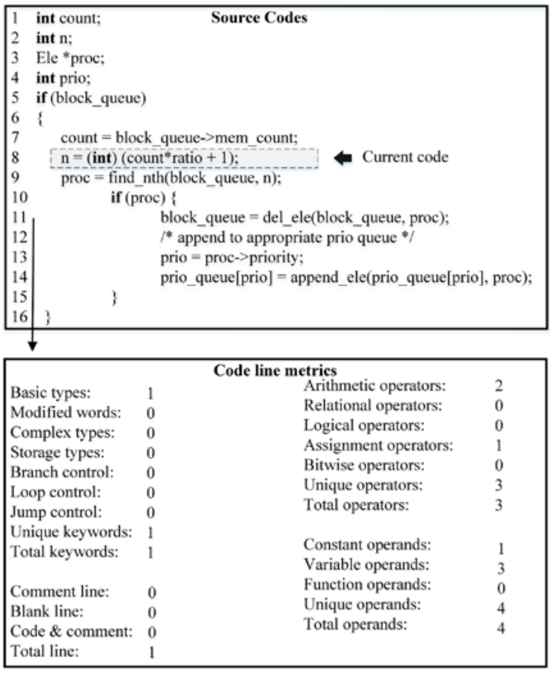


Fig. 2 The process of code feature extraction.

2.2.2. Feature extraction based on code statement

In the previous section, it was emphasized that the program code is composed of three vocabularies, that is, most of the defects in the code are basically related to them. For example, the wrong variable type, the wrong operator, and the wrong operand are defined respectively. Regardless of the defects that may be caused by the code running environment, the defects are highly related to these vocabularies. In this paper, keywords, operators, operands, and lines are mainly utilized as the metrics of each line of code for statement-based software defect prediction, as illustrated in Table 4.

Fig. 2 shows the process of extracting features from valid statements, which do not contain comment lines, empty lines, and brackets. The current code is line 8, including 1 basic type, 2 arithmetic operators, 3 variable operands, etc. The features and a defective class label together constitute the original dataset of software defect prediction.

2.2.3. Feature selection

It is a known fact that attribute selection has been extensively used in SDP [2, 16, 18, 24, 28, 45]. However, the following issues may arise during model training without using attribute selection (or feature selection): 1) The noise, redundancy and irrelevant features may be included. As all features collected in a dataset can be used for different tasks, not all of which contribute to SDP evaluation. 2) The accuracy, interpretability and generalization ability of the predictor may be affected.

Attribute selection can be primarily split into filter-based and wrapper-based approaches. The wrapper-based method obtains the best attribute subset by interacting with the learner feedback, while the filter-based method evaluates the feature subset on the training set without relying on the learner. Compared with the filter-based method, the wrapper-based method is inclined to more computationally complexity. Hence, a filter-based attribute selection method is preferred.

In this study, the classic correlation-based feature selection (CFS) filter method [14] is employed in software defect prediction [2, 3, 6].

Table 4. Simple description of code metrics

Features	Description
Keywords	Basic types, modified words, complex types, storage types, branch control, loop control, jump control, unique keywords, total keywords
Operators	Arithmetic operators, relational operators, logical operators, assignment operators, unique operators, total operators
Operands	Constant operands, variable operands, function operands, unique operands, total operands
Lines	Total line, comment line, blank line, code & comment
Defect	Bug

A subset of metrics is treated as input rather than all attributes of the dataset according to the heuristic function. This function (Equation 1) assigns a high score to a subset of metrics that are highly related to the class and have a low correlation with each other. Redundant metrics are therefore discriminated against because they are highly associated with one or more other metrics.

$$Merit_s = \frac{\overline{r_{cf}}}{\sqrt{k + k\bullet(k-1)\bullet\overline{r_{ff}}}} \tag{1}$$

Where Merits is the metric subset containing k metrics, $\overline{r_{cf}}$ is the average metric-class correlation, and $\overline{r_{ff}}$ is the average metric-metric intercorrelation.

2.2.4. Building the prediction model

The software defect prediction model based on valid code statements is a method that uses a predictor to find the relationship between features and the defect class label. Generally, software defect prediction models are trained and tested on the same project dataset [25, 28, 37, 39]. However, it is difficult to build the software defect prediction model when there is not sufficient historical data [44]. In order to avoid this problem, features and the defective label of the context scene are obtained by selecting similar tested software from other projects as training samples. In addition, metrics of the context scene is extracted from the tested software code as new data. By building and testing the statement-based prediction model in the training dataset and the new target testing dataset respectively, the probability of defect proneness of each line of code is calculated, as shown in Fig. 3.

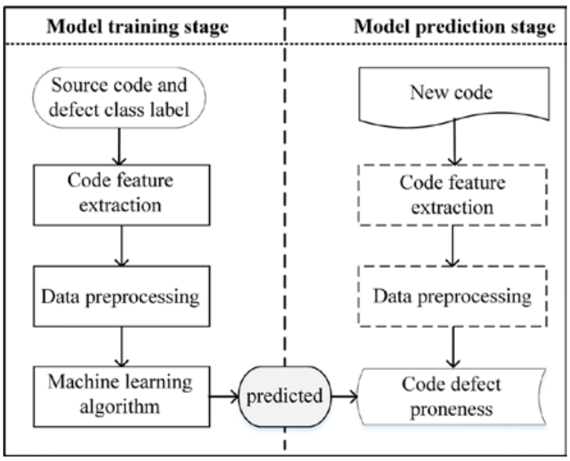


Fig. 3. The proposed software defect prediction framework based on code statementlevel

There is no difference between the defect prediction based on code context scenario and the traditional approach based on method/function or class defect prediction in the construction phase. The main difference lies in feature extraction that has been introduced in the previous part. There are plenty of machine learning methods, including supervised learning, unsupervised learning, and semi-supervised learning. Common supervised learning algorithms have been investigated for software reliability prediction and software defect prediction, such as Naive Bayes (NB) [2, 16, 28], linear Regression Model (LRM) [34], non-linear classifiers (e.g. support vector machines (SVM), neural network (NN)) [3, 49], ensemble learning (e.g. random forest (RF), bagging, Gradient Boost Regression Tree (GBRT)) [25, 47], and tree/rule-based classifiers (e.g. OneR, RIPPER, decision tree (C4.5), decision table (DT), partial decision trees (PART)) [28, 42]. Existing software defect predictors are mainly used for classification and ranking. Classification aims to predict software sample entities into defect proneness or non-defect proneness and the ranking task is to prioritize the samples according to the predicted code defect proneness probability. Compared with software defect classification, ranking results are more flexible and easier to use. Defect prediction can guide software testing to optimize resource allocation, which is essentially a test priority determination problem. This article uses code metrics for training to forecast the probability of defect proneness in code statements and then ranks them. Fig. 4 shows an example of the prediction results of the proposed method with the actual program code.

```
Lines|Probability|                        Source codes
7208|0.0175|    if (fastmap && startpos < total_size && !bufp->can_be_null)
7209|0.0000|  {
7210|0.0210|    if (range > 0)        /* Searching forwards.  */
7211|0.0000|      {
7212|0.0210|        register const char *d;
7213|0.0210|        register int lim = 0;
7214|0.0131|        int irange = range;
7215|0.0000|
7216|0.0061|            if (startpos < size1 && startpos + range >= size1)
7217|0.0025|              lim = range - (size1 - startpos);
7218|0.0000|
7219|0.0025|        d = (startpos >= size1 ? string2 - size1 : string1) + startpos;
7220|0.0000|
7221|0.0000|            /* Written out as an if-else to avoid testing `translate'
7222|0.0000|               inside the loop.  */
7223|0.0104|        if (translate)
7224|0.0324|              while (range > lim
7225|0.0025|                     && !fastmap[(unsigned char)
7226|0.0025|                             translate[(unsigned char) *d++]])
7227|0.0025|                range--;
7228|0.0025|        else
7229|0.0025|              while (range > lim && !fastmap[(unsigned char) *d++])
7230|0.0104|                range--;
7231|0.0000|
7232|0.0139|        startpos += irange - range;
7233|0.0000|      }
7234|0.0139|    else                          /* Searching backwards.  */
7235|0.0000|      {
7236|0.0139|        register char c = (size1 == 0 || startpos >= size1
7237|0.0139|                           ? string2[startpos - size1]
```

Fig. 4. An example of defect prediction result based on code statement-level

In Fig. 4, the leftmost column is the code line, the middle column is the defect prediction probability, and the right is the source code. We can see from this example that the software code statement-level defect prediction is feasible and has great research potential.

2.2. Test case Prioritization

The TCP problem is formalized as follows [35]:

Assumption *T*: A Test suite; *PT*: The set of permutations of *T*; *f*: a function from *PT* to the real numbers.

Problem: find $T' \in PT$ for $(\forall T'')(T'' \in PT)$ meets $f(T') \geq f(T'')$, where $(T'' \neq T')$.

In the formal description, PT represents all possible TCP suites. The input of function f is the specified priority order generated by the ranking target, and the output result is linked to the ranking target. Generally speaking, typical ranking targets include code coverage, defect detection rate, defect important, test cost and so on. In this paper, the code coverage-based TCP method concentrates on three aspects: prioritization strategy, prioritization criteria, and prioritization search.

Prioritization criteria describes the coverage criterion of the internal structure of the program code, such as method coverage, statement coverage, branch coverage and modified condition/decision coverage (MC/DC), which reflects the test adequacy. This paper associates code statement defect-proneness prediction with statement coverage to achieve test case sequencing. Prioritization search refers to the search algorithm used in sorting test cases such as greedy algorithm, genetic algorithm [26]. Most studies have adopted the greedy method, and empirical studies [11, 36] also prove that the greedy method is a simple and efficient ranking method. Prioritization strategy refers to the method used to prioritize test cases in a test suite. The typically used prioritization methods are as follows [11, 35]:

- Random prioritization: It assigns the priority of test cases in the test case set without any sorting criteria. It is used as a baseline for performance comparison in this paper.
- Optimal prioritization: It is based on the fact that each test case in the test suite can expose software errors, and determines the optimal sorting of test cases to maximize the defect detection rate. However, it is not a practical method because it requires pre-determining the defect information, which is often not available before testing. It was chosen to serve as an upper bound on the effectiveness of ranking strategies to distinguish the gaps between various strategies and the optimal solution.
- Total prioritization: It is a completely static strategy that directly counts the coverage of each test case in the test suite. The defect detection rate of test cases is calculated according to the code coverage and the test cases are sorted.
- Additional prioritization: It is a feedback mechanism strategy that takes into account overall coverage. Software entities (e.g. functions, statements, branches) covered by test cases are no longer considered. Therefore, after executing a test case, the covered software entities are eliminated and the remaining test cases are reordered. With the continuous execution of test cases, software entities will gradually be covered. When they are all covered, these entities need to be reset to the uncovered state, and the above process is repeated for the remaining test cases.
- Max prioritization: It takes the maximum defect prediction probability of the covered program code as the defect detection rate and the test cases are prioritized.

Prioritization strategy is usually combined with sorting criteria and sorting search. It is closely related to the average percentage faults detected (APFD) value, and different prioritization strategies often lead to different prioritization results. In this paper, we adopt a greedy search method based on code statement coverage to prioritize test cases on different strategies.

2.3. The defect detection rate of test cases

The proposed TCP-SCSDP determines the priority of test cases by their defect detection rate–a measure of how quickly defects are detected during software testing [11, 35]. Fig. 5 shows an example of defect detection rates for the total, additional and max prioritization strategies. It can be seen that the defect detection rate varies with the prioritization strategy.

Where the probability is obtained from software defect prediction based on code statement. The defect detection rate is computed according to the selected prioritization strategy and the code coverage of the test case. For example, the defect detection rate of test case 1 is 0.686 in total, 0.387 in max and 0.686 in additional, and the defect detection rate of test case 2 is 0.672 in total, 0.209 in max and 0.343 in additional prioritization. It is worth noting that the additional strategy is a feedback strategy, which will eliminate the covered code after the last test. After the defect detection rate of test cases is obtained and sorted, the APFD value is finally calculated.

Source codes	Probability	Test case 1	Test case 2
int main(int argc, char *argv[]	-		
{	-		
int x,y,z,m;	0.012	*	*
x=atoi(argv[1]);	0.015	*	*
y=atoi(argv[2]);	0.015	*	*
z=atoi(argv[3]);	0.015	*	*
if(y<z)	0.116	*	*
if(x<y)	0.116	*	*
m=y;	0.387	*	
else if(x<z)	0.174		*
m=y;	0.387		
else m=z;	0.209		*
else if(x>y)	0.174		
m=y;	0.387		
else if(x>z)	0.174		
m=x;	0.387		
else m=z;	0.209		
}	-		
Total		0.686	0672
Max		0.387	0.209
Additional		0.686	0.343

Fig. 5. Defect detection rate of test cases under different ranking strategies

3. Experimental setup

3.1. Research questions

In order to study the performance of the proposed method TCP-SCSDP systematically and objectively, we present three questions as follows:

RQ1: What is the performance comparison of different prioritization strategies?

To answer this question, we compared the APFD values under the five prioritization strategies of random, optimal, max, additional and total. Among them, the random strategy and the optimal strategy are used as a comparison baseline to evaluate the test performance of total, max and additional strategies.

RQ2: How do different software defect predictors affect test case prioritization?

The test prioritization task in this paper is the result of software defect prediction based on the code statement-level. Its test performance depends on the defect predictor. To answer this question, we choose LRM, GBRT and RT classifiers for comparative analysis.

3.2. Benchmark classifiers

For experimental comparison research, three distinct types of machine learning methods are used: Linear Regression Model (LRM), Random Forest (RF), and Gradient Boosting Regression Tree (GBRT). The three classifiers are widely used in software defect prediction and show superior prediction performance [25, 34, 47]. In addition, because this paper is a test case prioritization based on code statement-level defect prediction, all three algorithms can well support the defect prediction.

LRM: A curve that is called the best fitting curve is utilized to fit the data points, and the fitting process is called regression. When the curve is linear, the process is called the linear regression. The main idea of linear regress model is to use the pre-determined weights to combine the attributes to represent the categories.

$$x = w_0 + w_1 a_1 + w_2 a_2 + \cdots + w_k a_k \tag{2}$$

where x is a class label, $a_1, a_2, \cdots, a_k$ is an attribute value, $w_1, w_2, \cdots, w_k$ is a weight value. In general, LRM can be resolved by the least square method.

RF: It is an improved decision tree algorithm, which is a typical bagging ensemble learning, mainly to handle the over-fitting problem of decision tree. It has the advantages of high accuracy, not easy to overfit and can handle high-dimensional data [47]. It uses multiple the mechanism of decision trees for voting to improve the prediction results. Assume that random forest is composed of m trees, where each tree is generated by a certain amount of training samples n. In order to ensure the generalization ability of random forest, n samples are generated by bootstrapping and the final prediction results are obtained by bagging.

GBRT: It is a gradient boosting algorithm, which was first put forward by Friedman [12]. It fits a regression tree by using the fastest descent approximation method, that is, using the value of the negative gradient of the loss function in the current model as the approximate value (pseudo-residual) of the residual of the lifting tree algorithm in the regression problem. In simple terms, each tree in a progressive gradient regression tree is learned from the residuals of all previous trees.

3.3. Benchmark datasets

In this paper, four open-source C program datasets are chosen as experimental objects, which are *Gzip*, *Grep*, *Flex* and *Sed*. Among them, *Gzip* is a widely-used file compression program of GNU free software, *Grep* is a text search tool running under Linux which can search text using specific pattern matching including regular expressions, *Flex* is a program for SQL lexical analysis in Linux environment and sed is a tool for running Linux instructions. These programs are intensively studied in the field of software engineering [4, 17]. Their source code and related materials can be accessed from SIR. According to the defect information found in the historical version of these codes, defect injection is carried out selectively, in which the defect of the deleted class and non-modified code class are not injected. Table 5 describes the basic information of the datasets.

Table 5. An overview of subjects used in this study

Dataset	Size	Injected defects	Defect rate per thousand lines of code (KLOC)	Number of test cases	Source	Description
Gzip	5680	37	6.51	279	SIR	File compression utility
Grep	10068	47	4.67	669	SIR	Text search tool
Flex	10459	32	3.06	447	SIR	SQL parsing tool
Sed	14427	27	1.87	261	SIR	Linux command run tool

3.4. Performance evaluation measures

The goal of test case prioritization is to find as many software defects as early as possible, so as to they can be fixed early in testing. It can improve the effectiveness of software testing and shorten the software development life cycle. Generally, APFD is invoked as the performance evaluation indicator of the priority ranking method. Suppose there are n test cases in a test case T, m defects found in a defective set F, and the test case rank T', its APFD is as follows:

$$APFD = 1 - \frac{TF_1 + TF_2 + \cdots + TF_i + \cdots TF_m}{m * n} + \frac{1}{2n} \tag{3}$$

In Equation 3, TF_i is the position of the first test case with defect i found in sorting strategy T'. The APFD value ranges from 0 to 1. The higher its value is, the better the test case prioritization is. For instance, there is a test case set as shown in Table 6. If it is sorted according to $T_1 - T_{10}$, the APFD value of the test sequence is $APFD = 0.46$. If it is sorted according to $T_{10} - T_1$, the APFD value of the test sequence is $APFD = 0.7875$.

Table 6. Correspondence between test cases and defects.

Test Case ID	Defect ID							
	F1	F2	F3	F4	F5	F6	F7	F8
T1								
T2								
T3		※					※	
T4	※		※				※	
T5		※					※	
T6								
T7		※	※	※	※			
T8	※		※					
T9						※		
T10		※		※				※

4. Experimental results and discussion

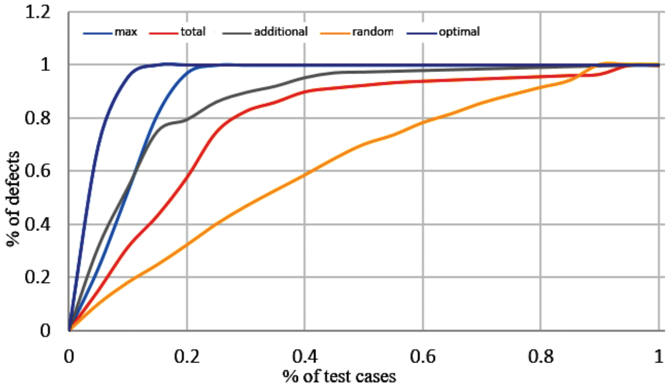
In order to systematically study the problems raised, four open-source program datasets of SIR website are taken as experimental objects. Taking the set of Siemens programs (*Gzip*, *Grep*, *Flex* and *Sed*) as the data source, data samples are established through the automatically extracted code features and defect label as training data-set of the prediction model. And the testing dataset is removed from the training dataset to avoid over-fitting seen in Table 7. A prediction model trained from the historical data across other projects is utilized to predict defects in the project to solve the problem of insufficient historical defect data. For instance, assuming that the training dataset is {*Grep*, *Flex*, *Sed*} then the *Gzip* program is the testing dataset, and then the *Gzip* program is the testing dataset.

Aiming at the first problem RQ1, five test case prioritization strategies, max, total, additional, random and optimal, are compared. Because the random strategy has randomness in selecting test cases, the average result of 20 times will be used as a comparison value to objectively evaluate its performance. In addition, for the second problem RQ2, three different predictors, LRM, RF, and GBRT,

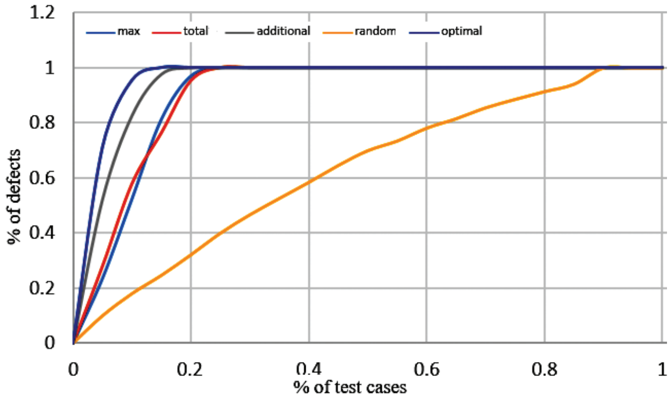
Table 7. Case studies

Training	Testing (case study)
Grep, Flex, Sed	Gzip
Gzip, Flex, Sed	Grep
Gzip, Grep, Sed	Flex
Gzip, Grep, Flex	Sed

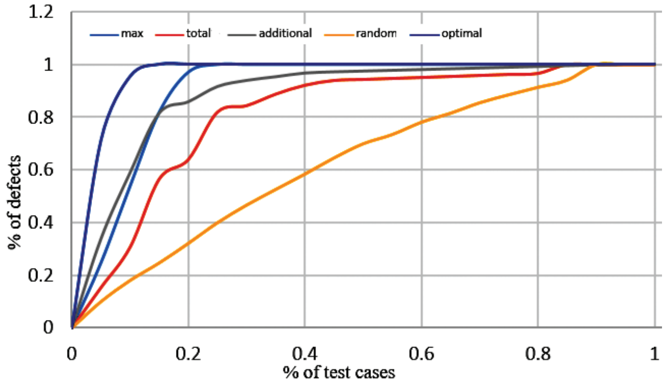
are used to analyse the impact of on the ranking of test cases. Fig. 6-9 show the Alberg diagrams [31, 34] of the experimental results, where the x-axis indicates the proportion of test cases used in the total test cases and the y-axis indicates the percentage of defects found in the total defects.



(a) LRM

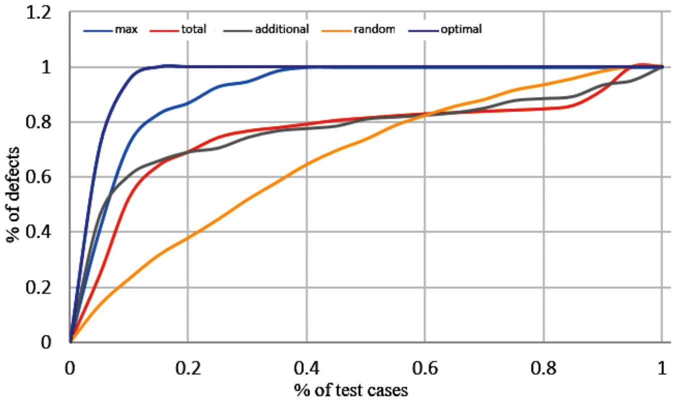


(b) RF

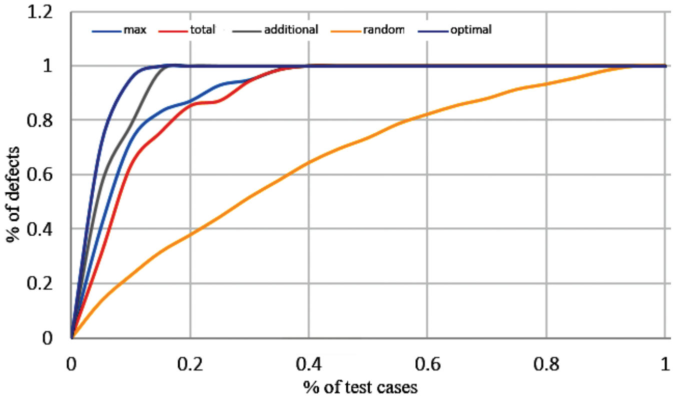


(c) GBRT

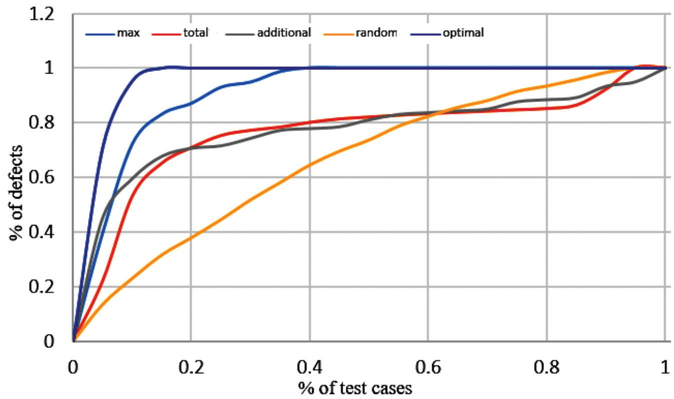
Fig. 6. Results on Flex dataset from different combinations of predictors and test prioritization strategies



(a) LRM



(b) RF



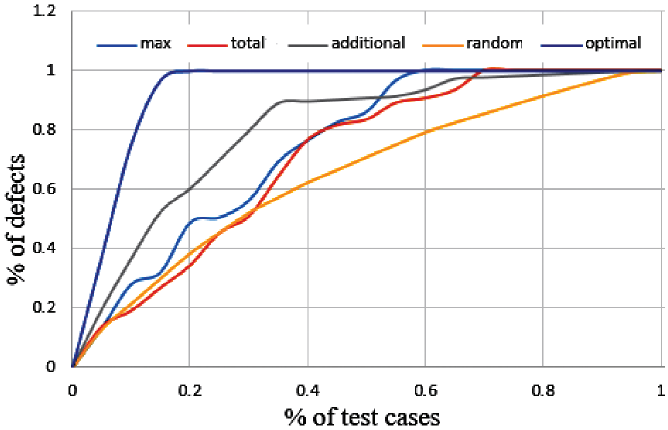
(c) GBRT

Fig. 7. Results on Grep dataset from different combinations of predictors and test prioritization strategies

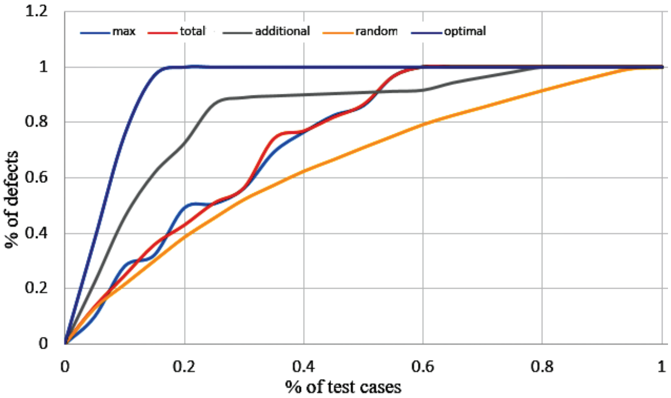
4.1. RQ1: What is the performance comparison of different prioritization strategies?

From the observation of Fig. 6-9 and Table 8, it can be seen that the performance of different prioritization strategies is different.

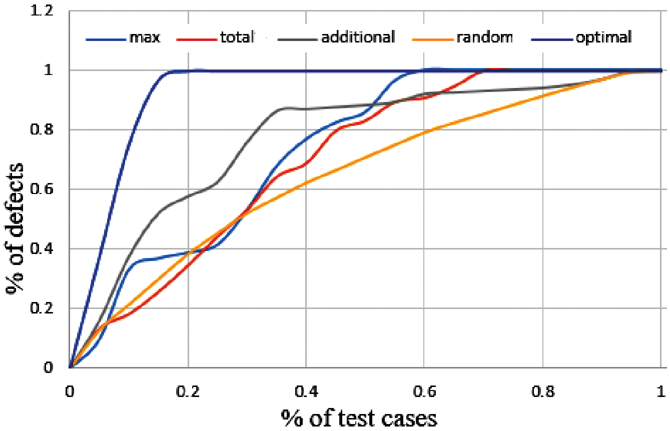
- The optimal strategy has the best performance, which is the upper theoretical limit of test case prioritization. The APFD values of the optimal strategy in *Gzip* and *Sed* are 0.933 and 0.948, respectively.
- The random strategy has the worst performance, which is caused by its indiscriminate random selection of test cases. Its APFD values in *Gzip*, *Grep*, *Flex* and *Sed* are 0.628, 0.65, 0.62 and 0.551, respectively, which are close to the theoretical values of 0.5. However, the actual result of the random strategy is higher than the theoretical value. This is because the corresponding relationship between the designed test cases and defects is not one-to-one, that is, a defect can be found by multiple test cases, which is closer to the actual test situation. This leads to an increased probability of finding defects in the test case set. Therefore, the probability of finding defects in the test case is higher than the theoretical value.
- The curve of the max strategy on the *Gzip* and *Sed* datasets is closer to that of the total (see Fig. 8 and Fig. 9), and it is closer to the additional strategy on the *Grep* and *Flex* datasets (see



(a) LRM



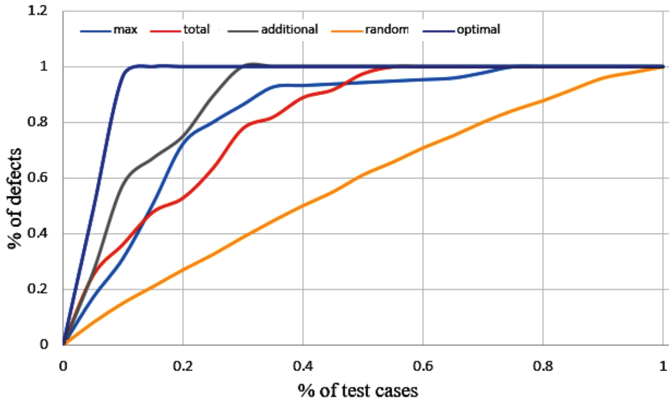
(b) RF



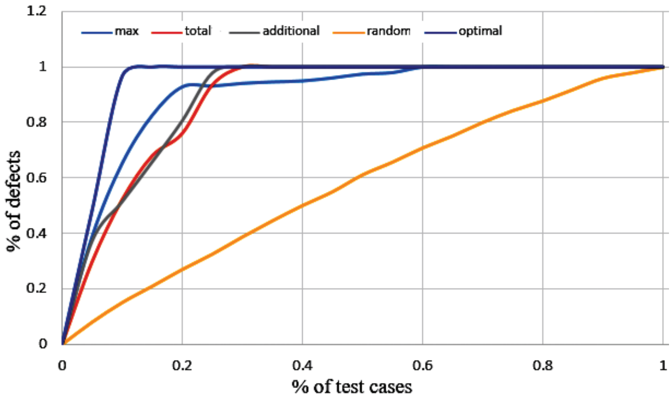
(c) GBRT

Fig. 8. Results on Gzip dataset from different combinations of predictors and test prioritization strategies

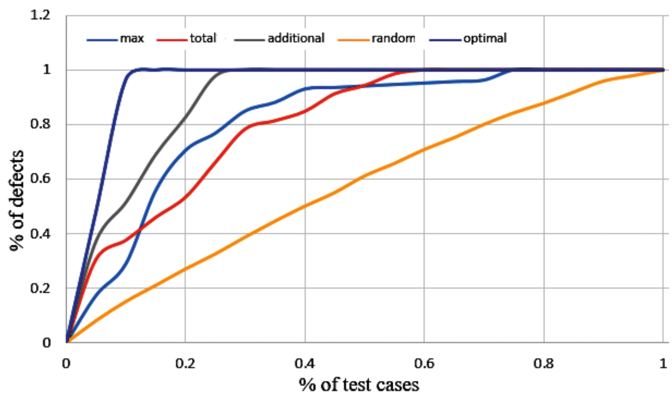
Fig. 7 and Fig. 6). The APFD values of the max strategy are generally higher than those of the total strategy (see Table 8). However, the starting curve of the total is higher than the max, which indicates that the total is actually better than the max in the early stage of testing. This is mainly because the total is on the overall coverage strategy. It is easier to prioritize test cases with wider coverage in the early stage so that the probability of finding defects is higher. However, it is prone to falling into the situation of repeatedly covering the tested code later in the testing. The max strategy takes into account the greatest probability of detecting defects in the code. Although it is not as good as the



(a) LRM



(b) RF



(c) GBRT

Fig. 9. Results on Sed dataset from different combinations of predictors and test prioritization strategies

total strategy in the early stage of testing, it is easier to prioritize the execution of defective test cases in the later stage.

- The additional is a strategy with excellent performance, which is only slightly lower than the optimal strategy. Although the additional is not as good as the max on *Grep* and *Flex* datasets in some cases, the overall curve in the figures shows that the additional is worse than the max and total. This is because both the max and total are static strategies, while the additional is a dynamic one. After each test execution, the defect detection rate of the test case will be readjusted according to the coverage of the test code. This feedback method is helpful for the additional strategy to optimize its test case selection behaviour in real-time.
- In addition, compared with the random strategy, the test performance gain curves of the three TCP strategies max, total and

optimal are convexly shown in Fig. 10. As the proportion of test cases increases, the test performance difference tends to increase firstly and then decreased. Moreover, when 20%-30% of the test cases are selected, three test performance gain curves reach the maximum and the efficiency ratio is the largest.

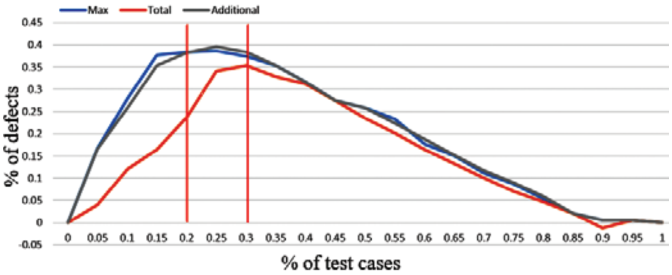


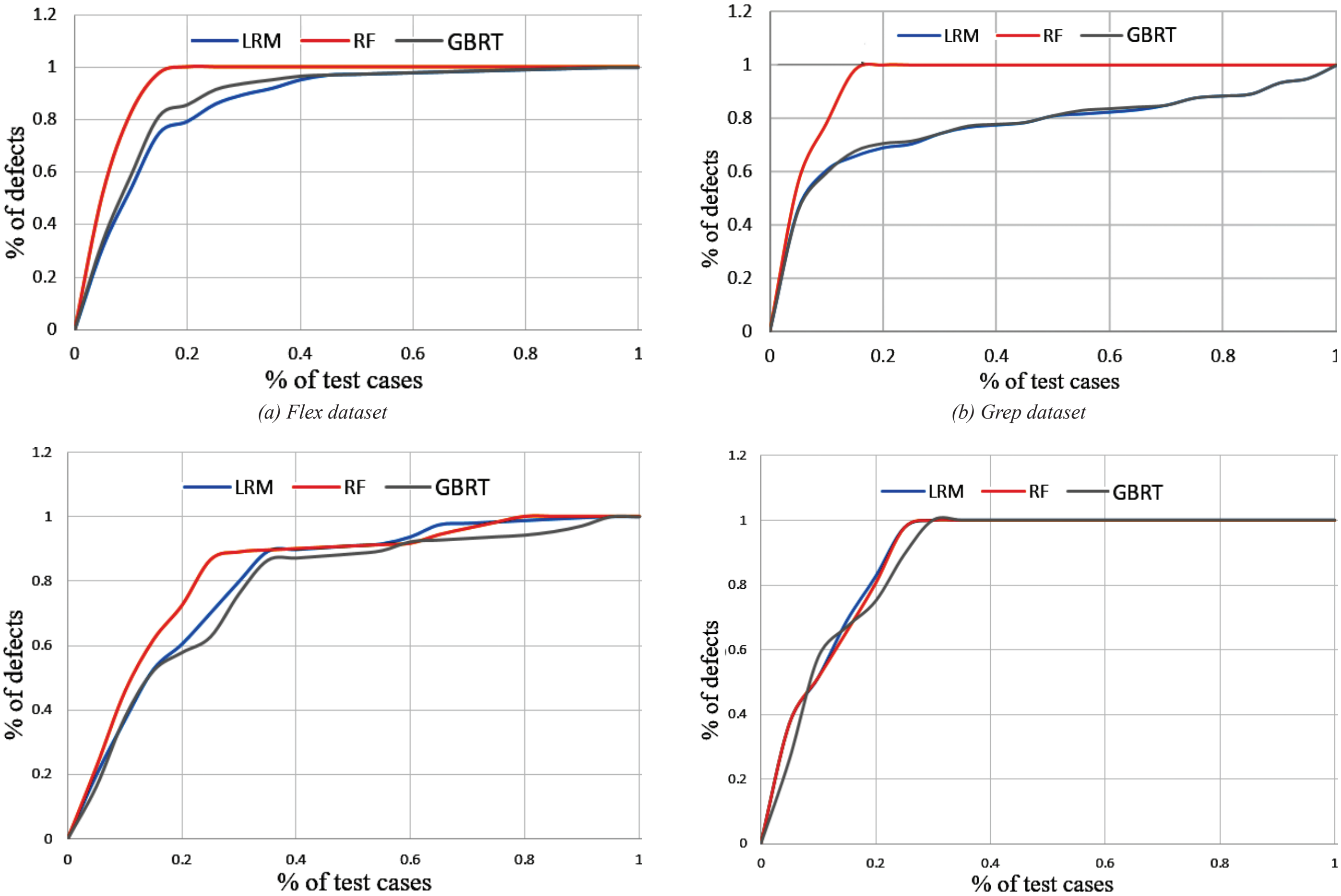
Fig. 10. The test performance gains of the three strategies max, total and optimal compared to the random strategy

Based on the above results, the following conclusions can be drawn: The proposed model is suitable for test case prioritization. In the test case prioritization methods, the APFD performance of the additional strategy is preferable to max strategy and total strategy.

5.2. RQ2: How do different software defect predictors affect test case prioritization?

This problem explores the impact of different defect prediction classifiers on test results. It can be found that from Table 8 that different predictors have an impact on the additional, max and total strategies, while the optimal and random strategies will not be disturbed because they are not related to the predictor. For the convenience of analysis, this section analyses the performance of the three predictors on *Gzip*, *Grep*, *Flex* and *Sed* by taking the additional strategy as the object as illustrated in Figure 11.

It can be seen from Fig. 11 that using the RF-based predictor for the additional test case prioritization strategy is better than using



(a) Flex dataset *(b) Grep dataset*

(c) Gzip dataset *(d) Sed dataset*

Fig. 11. Impact of LRM, RF and GBRT predictors on the additional prioritization strategy

On the whole, the performance of the test case prioritization strategy is optimal, additional, max, total and random in turn from high to low. In order to quantitatively analyse the performance difference, the APFD average values under different classifiers and test case prioritization strategies are calculated as shown in Table 8. The average value is used as total, max, additional and optimal strategies are 0.7918 (-16.05%), 0.8451 (-10.72%), 0.8473 (-10.50%), and 0.9523 respectively. The percentiles in parentheses are the performance loss ratio relative to the optimal strategy. The smaller the loss ratio, the better the ranking.

the LRM-based or GBRT-based on *Gzip*, *Grep*, *Flex* and *Sed* datasets. Especially on the *Grep* dataset, the test performance is significantly higher than the other two predictors. This shows that the choice of the predictor has an influence on the test case prioritization based on code statement level defect prediction, and its degree of influence is also inconsistent with the different experimental objects.

Based on the above results, the following conclusions can be drawn: The prediction performance of the defect predictor will affect the efficiency of the test case prioritization method, that is, different

Table 8. APFD values under different prioritization and predictors

Dataset	prioritization	Machine learning algorithm		
		LRM	RF	GBRT
Gzip	Max	0.740	0.739	0.731
	Total	0.701	0.740	0.698
	Additional	0.794	0.821	0.766
	Random	0.628	0.628	0.628
	Optimal	0.933	0.933	0.933
Grep	Max	0.910	0.910	0.908
	Total	0.746	0.894	0.751
	Additional	0.746	0.941	0.766
	Random	0.650	0.650	0.650
	Optimal	0.964	0.964	0.964
Flex	Max	0.900	0.900	0.908
	Total	0.776	0.907	0.802
	Additional	0.855	0.943	0.874
	Random	0.620	0.620	0.620
	Optimal	0.964	0.964	0.964
Sed	Max	0.814	0.873	0.808
	Total	0.801	0.886	0.799
	Additional	0.880	0.889	0.892
	Random	0.551	0.551	0.551
	Optimal	0.948	0.948	0.948

defect predictors will bring different prediction probabilities to the test case prioritization.

5. Threats to validity

As an empirical study, its potential limitations must be taken into account when interpreting its results. This section describes several potential threats to the validity of our models.

5.1. Internal Validity

Internal validity mainly refers to the correctness and reproducibility of our empirical results. We implement all the baseline predictors (RF, LRM and GBRT) with default settings invoking WEKA to reduce the potential possibility to make mistakes. The optimal parameters of the predictors may be different for different defect datasets, which may lead to different results. However, it does not hinder the feasibility and effectiveness of the test case prioritization based on code statement-level software defect prediction. And all the algorithms involved adopting the same data pre-processing (e.g. CFS-based attribute selection [14]) to minimize redundancy. Besides, we have double-checked all of our experiments, but there may be a few errors.

5.2. External Validity

External validity relates to the generalization ability of our empirical results. The proposed approach was compared and analysed in four selected subjects, which may have data quality issues. And the attributes collected are all code statement metrics from SIR Repository and the samples are abstracted at code statement level from C programming language. Although our proposed model could be employed in other programming languages (i.e., Java, C++), we cannot guarantee the same empirical results. Besides, we only employed the three prevailing defect classifiers (RF, LRM, and GBRT). As we all know, there are a large number of predictive classifiers [6, 25, 47]. We could not validate all other algorithms due to time and space constraints. However, it does not dispute that choosing different predictors affects test case prioritization results. To reduce the external threats, more programming languages, high-quality defect datasets, and the predictors should be utilized in the future.

5.3. Construct Validity

Construct validity refers to the suitability of the test performance evaluation measure. There are some several measures [7], such as average severity of faults detected (ASFD), coverage effectiveness (CE), total percentage of faults detected (TPFD), average percentage faults detected (APFD). However, in fact, there is no studies have applied all of the measures to evaluate test case prioritization. We chose carefully the most commonly used measure APFD to prioritize test cases. Besides, in order to reduce construct validity, we also use the Alberg diagram to visually describe the curve of the proportion of found defects to the total number of defects as the proportion of test cases to the total number of test cases increases.

6. Conclusion

This paper proposes a novel test case prioritization method based on code statement-level defect prediction named TCP-SCSDP, which takes into account the possible distribution of defects and prediction granularity. The proposed feature set for measuring code statements first is used as the input for the statement software defect prediction model, and data pre-processing is performed on the software defect dataset. Secondly, the predictor is applied to predict the defect proneness probability of valid code statements. Then the defect detection rate of all test cases is calculated by using the test case prioritization strategy, and they are sorted from high to low. Finally, APFD is used to evaluate the prioritization.

Experimental results on 4 open source datasets show that the proposed approach is feasible and effective, and the test performance will be affected by the predictor and the test case prioritization strategy.

Our future work will focus on the following aspects: (1) Collect more open source software projects, programming languages (e.g. C++, Java) and predictors (e.g. neural network, k nearest neighbour), as mentioned earlier, to validate the generality of our method. (2) Optimize the test case prioritization strategy to improve software testing efficiency.

Acknowledgement

The research work is supported by a grant from the Science & Technology on Reliability & Environmental Engineering Laboratory of China (Grant No.614200404031117). And this work is also partially supported by No. 61400020404. We would be grateful to the editor and anonymous reviewers for their insightful and valuable comments and suggestions to improve the paper.

References

1. Arafeen M J, Do H. Test case prioritization using requirements-based clustering. Proceedings of the 2013 IEEE Sixth International Conference on Software Testing, Verification and Validation 2013: 312-321, https://doi.org/10.1109/ICST.2013.12.
2. Arar O F, Ayan K. A feature dependent naive bayes approach and its application to the software defect prediction problem. Applied Soft Computing 2017; 59: 197-209, https://doi.org/10.1016/j.asoc.2017.05.043.

3. Arar O F, Ayan K. Software defect prediction using cost-sensitive neural network. Applied Soft Computing 2015; 33(C): 263-277, https://doi.org/10.1016/j.asoc.2015.04.045.
4. Bertolino A, Miranda B, Pietrantuono R, et al. Adaptive coverage and operational profile-based testing for reliability improvement. Proceedings of the 39th International Conference on Software Engineering 2017: 541-551.
5. Boehm B, Basili V R. Software defect reduction top 10 list. Computer 2001; 34(1): 135-137, https://doi.org/10.1109/2.962984.
6. Catal C, Diri B. Investigating the effect of dataset size, metrics sets, and feature selection techniques on software fault prediction problem. Information Sciences 2009; 179(8): 1040-1058, https://doi.org/10.1016/j.ins.2008.12.001.
7. Catal C, Mishra D. Test case prioritization: A systematic mapping study. Software Quality Journal 2013; 21(3): 445-478, https://doi.org/10.1007/s11219-012-9181-z.
8. Di Nardo D, Alshahwan N, Briand L, et al. Coverage-based regression test case selection, minimization and prioritization: A case study on an industrial system. Software Testing, Verification and Reliability 2015; 25(4): 371-396, https://doi.org/10.1002/stvr.1572.
9. Do H, Mirarab S, Tahvildari L, et al. The effects of time constraints on test case prioritization: A series of controlled experiments. IEEE Transactions on Software Engineering 2010; 36(5): 593-617, https://doi.org/10.1109/TSE.2010.58.
10. Durelli V H S, Durelli R S, Borges S S, et al. Machine learning applied to software testing: A systematic mapping study. IEEE Transactions on Reliability 2019; 68(3): 1189-1212, https://doi.org/10.1109/TR.2019.2892517.
11. Elbaum S, Malishevsky A G, Rothermel G. Test case prioritization: A family of empirical studies. IEEE Transactions on Software Engineering 2002; 28(2): 159-182, https://doi.org/10.1109/32.988497.
12. Friedman J H. Stochastic gradient boosting. Computational statistics & data analysis 2002; 38(4): 367-378.
13. Garg D, Datta A. Test case prioritization due to database changes in web applications. Proceedings of the 2012 IEEE Fifth International Conference on Software Testing, Verification and Validation 2012: 726-730.
14. Hall M, Frank E, Holmes G, et al. The weka data mining software: An update. ACM SIGKDD explorations newsletter 2009; 11(1): 10-18.
15. Harrold M J. Testing: A roadmap. Proceedings of the Future of Software Engineering 2000: 61-72.
16. He P, Li B, Liu X, et al. An empirical study on software defect prediction with a simplified metric set. Information and Software Technology 2015; 59(C): 170-190, https://doi.org/10.1016/j.infsof.2014.11.006.
17. Henard C, Papadakis M, Harman M, et al. Comparing white-box and black-box test prioritization. Proceedings of the 2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE) 2016: 523-534, https://doi.org/10.1145/2884781.2884791.
18. Hosseini S, Turhan B, Mäntylä M. A benchmark study on the effectiveness of search-based data selection and feature selection for cross project defect prediction. Information and Software Technology 2017; 95: 296-312, https://doi.org/10.1016/j.infsof.2017.06.004.
19. Hovemeyer D, Pugh W. Finding bugs is easy. Acm sigplan notices 2004; 39(12): 92-106, https://doi.org/10.1145/1052883.1052895.
20. Huang C Y, Chang J R, Chang Y H. Design and analysis of gui test-case prioritization using weight-based methods. Journal of Systems and Software 2010; 83(4): 646-659, https://doi.org/10.1016/j.jss.2009.11.703.
21. Khatibsyarbini M, Isa M A, Jawawi D N A, et al. Test case prioritization approaches in regression testing: A systematic literature review. Information and Software Technology 2018; 93: 74-93, https://doi.org/ 10.1016/j.infsof.2017.08.014.
22. Krishnamoorthi R, Sahaaya Arul Mary S A. Requirement based system test case prioritization of new and regression test cases. International Journal of Software Engineering and Knowledge Engineering 2009; 19(3): 453-475, https://doi.org/10.1142/S0218194009004222.
23. Lachmann R, Schulze S, Nieke M, et al. System-level test case prioritization using machine learning. Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA) 2016: 361-368, https://doi.org/10.1109/ICMLA.2016.163.
24. Laradji I H, Alshayeb M, Ghouti L. Software defect prediction using ensemble learning on selected features. Information and Software Technology 2015; 58: 388-402, https://doi.org/10.1016/j.infsof.2014.07.005.
25. Lessmann S, Baesens B, Mues C, et al. Benchmarking classification models for software defect prediction: A proposed framework and novel findings. IEEE Transactions on Software Engineering 2008; 34(4): 485-496, https://doi.org/10.1109/TSE.2008.35.
26. Li Z, Harman M, Hierons R M. Search algorithms for regression test case prioritization. IEEE Transactions on Software Engineering 2007; 33(4): 225-237, https://doi.org/10.1109/TSE.2007.38.
27. Mei H, Hao D, Zhang L, et al. A static approach to prioritizing junit test cases. IEEE Transactions on Software Engineering 2012; 38(6): 1258-1275, https://doi.org/10.1109/TSE.2011.106.
28. Menzies T, Greenwald J, Frank A. Data mining static code attributes to learn defect predictors. IEEE Transactions on Software Engineering 2006; 33(1): 2-13, https://doi.org/10.1109/TSE.2007.256941.
29. Mirarab S, Tahvildari L. A prioritization approach for software test cases based on bayesian networks. Proceedings of the International Conference on Fundamental Approaches to Software Engineering 2007: 279-290.
30. Nam J, Kim S. Clami: Defect prediction on unlabeled datasets. Proceedings of the 2015 30th IEEE/ACM International Conference on Automated Software Engineering (ASE) 2015: 452-463, https://doi.org/10.1109/ASE.2015.56.
31. Ohlsson N, Alberg H. Predicting fault-prone software modules in telephone switches. IEEE Transactions on Software Engineering 1996; 22(12): 886-894, https://doi.org/10.1109/32.553637.
32. Paterson D, Campos J, Abreu R, et al. An empirical study on the use of defect prediction for test case prioritization. Proceedings of the 2019 12th IEEE Conference on Software Testing, Validation and Verification (ICST) 2019: 346-357, https://doi.org/10.1109/ICST.2019.00041.
33. Peng R, Zhai Q. Modeling of software fault detection and correction processes with fault dependency. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19(3): 467-475, https://doi.org/10.17531/ein.2017.3.18.
34. Peng X, Liu B, Wang S. Feedback-based integrated prediction: Defect prediction based on feedback from software testing process. Journal of Systems and Software 2018; 143: 159–171, https://doi.org/10.1016/j.jss.2018.05.029.
35. Rothermel G, Untch R H, Chu C, et al. Prioritizing test cases for regression testing. IEEE Transactions on Software Engineering 2001; 27(10): 929-948, https://doi.org/10.1109/32.962562.
36. Rothermel G, Untch R H, Chu C, et al. Test case prioritization: An empirical study. Proceedings of the IEEE International Conference on Software Maintenance-1999 (ICSM'99)'Software Maintenance for Business Change'(Cat No 99CB36360) 1999: 179-188.
37. Shao Y, Liu B, Wang S, et al. A novel software defect prediction based on atomic class-association rule mining. Expert Systems with Applications 2018; 114: 237-254, https://doi.org/10.1016/j.eswa.2018.07.042.

38. Shatnawi O. Measuring commercial software operational reliability: An interdisciplinary modelling approach. Eksploatacja i Niezawodnosc – Maintenance and Reliability, 2014; 16(4): 585-594.
39. Song Q, Jia Z, Shepperd M, et al. A general software defect-proneness prediction framework. IEEE Transactions on Software Engineering 2011; 37(3): 356-370, https://doi.org/10.1109/TSE.2010.90.
40. Spieker H, Gotlieb A, Marijan D, et al. Reinforcement learning for automatic test case prioritization and selection in continuous integration. Proceedings of the 26th ACM SIGSOFT International Symposium on Software Testing and Analysis 2017: 12-22, https://doi.org/10.1145/3092703.3092709.
41. Srikanth H, Hettiarachchi C, Do H. Requirements based test prioritization using risk factors: An industrial study. Information and Software Technology 2016; 69: 71-83, https://doi.org/10.1016/j.infsof.2015.09.002.
42. Sun Z, Song Q, Zhu X. Using coding-based ensemble learning to improve software defect prediction. IEEE Transactions on Systems Man and Cybernetics Part C 2012; 42(6): 1806-1817, https://doi.org/10.1109/TSMCC.2012.2226152.
43. Tonella P, Avesani P, Susi A. Using the case-based ranking methodology for test case prioritization. Proceedings of the 2006 22nd IEEE International Conference on Software Maintenance 2006: 123-133.
44. Turhan B, Menzies T, Bener A c, ,e B., et al. On the relative value of cross-company and within-company data for defect prediction. Empirical Software Engineering 2009; 14(5): 540-578, https://doi.org/10.1007/s10664-008-9103-7.
45. Wang H, Khoshgoftaar T M, Napolitano A. Software measurement data reduction using ensemble techniques. Neurocomputing 2012; 92(3): 124-132, https://doi.org/10.1016/j.neucom.2011.08.040.
46. Wang S, Nam J, Tan L. Qtep: Quality-aware test case prioritization. Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering 2017: 523-534, https://doi.org/10.1145/3106237.3106258.
47. Wang S, Yao X. Using class imbalance learning for software defect prediction. IEEE Transactions on Reliability 2013; 62(2): 434-443, https://doi.org/10.1109/TR.2013.2259203.
48. Xiao L, Miao H, Zhuang W, et al. An empirical study on clustering approach combining fault prediction for test case prioritization. Proceedings of the 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS) 2017: 815-820.
49. Xu J, Meng Z, Xu L. Integrated system of health management-oriented reliability prediction for a spacecraft software system with an adaptive genetic algorithm support vector machine. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16(4): 571-578.
50. Yang X, Tang K, Yao X. A learning-to-rank approach to software defect prediction. IEEE Transactions on Reliability 2015; 64(1): 234-246, https://doi.org/10.1109/TR.2014.2370891.
51. Yoo S, Harman M. Regression testing minimization, selection and prioritization: A survey. Software Testing, Verification and Reliability 2012; 22(2): 67-120, https://doi.org/10.1002/stv.430.

Yuanxun SHAO
Bin LIU
Shihai WANG
School of Reliability and Systems Engineering
Beihang University
No.37 Xueyuan RD. Haidian, 100191, Beijing, China

Peng XIAO
Ji'an Municipal Industry and Information Technology Bureau
11/F, Block B, Administration Center Building
Jizhou District, Ji'an, Jiangxi, 343000, China

E-mails: yuanxunshao@buaa.edu.cn, liubin@buaa.edu.cn, wangshihai@buaa.edu.cn, buaaxp@foxmail.com

Article citation info:
DZIERWA A, GAŁDA L, TUPAJ M, DUDEK K. Investigation of wear resistance of selected materials after slide burnishing process. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 432–439, http://dx.doi.org/10.17531/ein.2020.3.5.

Andrzej DZIERWA
Lidia GAŁDA
Mirosław TUPAJ
Kazimiera DUDEK

INVESTIGATION OF WEAR RESISTANCE OF SELECTED MATERIALS AFTER SLIDE BURNISHING PROCESS

BADANIA ODPORNOŚCI NA ZUŻYCIE WYBRANYCH MATERIAŁÓW PODDANYCH PROCESOWI NAGNIATANIA ŚLIZGOWEGO*

The article presents the research on the impact of slide burnishing process carried out with use of various ceramics on friction and wear of steel elements. In addition, surfaces after grinding, lapping and polishing processes were tested. The tribological couple was made of steel discs, toughened to a hardness of 40 ± 2 HRC, and balls made of 100Cr6 steel with a hardness of 62 HRC. The tests were carried out at three sliding speeds: 0.16 m/s, 0.32 m / s and 0.48 m/s. The research proved the possibility of improving selected tribological properties of friction pairs thanks to the use of slide burnishing process and also allowed to establish a number of relationships between the parameters characterizing the surface topography and the tribological parameters.

Keywords: *surface topography, friction, wear.*

W artykule przedstawiono wyniki badań wpływu procesu nagniatania ślizgowego realizowanego z wykorzystaniem różnych ceramik na wielkość zużycia oraz siłę tarcia elementów stalowych. Dodatkowo badaniom poddano powierzchnie po procesach szlifowania, docierania oraz polerowania. Skojarzenie materiałowe stanowiły tarcze stalowe ulepszone cieplnie do twardości 40±2 HRC oraz kulki ze stali100Cr6 o twardości 62 HRC. Badania zrealizowano przy trzech prędkościach poślizgu: 0,16 m/s, 0,32 m/s oraz 0,48 m/s. Badania udowodniły możliwość poprawy wybranych właściwości tribologicznych par trących dzięki zastosowaniu procesu nagniatania ślizgowego a także pozwoliły na ustalenie szeregu zależności pomiędzy parametrami charakteryzującymi strukturę geometryczną powierzchni oraz parametrami tribologicznymi.

Słowa kluczowe: *struktura geometryczna powierzchni, tarcie, zużycie.*

List of abbreviations and symbols:

F_{30} – value of the friction force obtained for the sliding distance equal to 30 m,
$F_{śr}$ – average value of the friction force,
Hm – maximum values of the microhardness of the surface layer,
σ_{max} – maximum values of residual stresses of the surface layer,
Sal – fastest-decay autocorrelation length,
Sk – core roughness depth,
Sku – kurtosis,
Spd – peak density,
Spk – reduced peak height,
Sq – root mean square height of the surface,
Ssk – skewnees,
Svk – reduced valley depth,
Sz – maximum height of the surface,
VD – wear volume of the discs,

1. Introduction

Any new technical object constructed in accordance with the requirements contained in the design and technological documentation has its own, full operational potential. During operation, and in the result of work performed this potential is decreased, in the result of the occurring physicochemical changes of elements, i.e. the wear of friction pairs, material fatigue, corrosive processes, etc. [13]. These changes can reduce reliability, increase failure rate, or reduce object performance. Destruction processes occur in the result of working conditions, referred to as wear [36], which cause a sudden or gradual loss of performance of the elements. Because wear in most cases leads to a reduction in the operating potential of machines and their components it should be countered. This countermeasures should start at the design stage, with appropriate selection of elements of the tribomechanical assembly, so as to reduce wear during operation. In addition to design, a number of technological methods of wear prevention are also applied. These can include, among others, the following [1, 4, 38]:

- use of heat and thermo-chemical treatment (e.g. harden nitriding, carburizing, cyaniding),
- use of plastic forming of metals (e.g. hammering, burnishing),
- applying overlays and coatings (e.g. chemical nickel plating, hard facing).

Burnishing is one of the surface treatments that can delay wear processes [22]. It includes ways such as roller burnishing [11], roller finishing, slide burnishing [17] and even similar techniques like shotpeening [9, 16]. In the case of slide burnishing, a hard and smooth burnishing element is pressed against the machined surface with the appropriate force, causing sliding friction in the burnishing zone and, as in the consequence of this process, smoothing of the surface and beneficial changes in the properties of the surface layer of the object

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

[31]. By using slide burnishing, we can improve hardness of the [28, 33], attain good surface smoothness [14, 25, 30], form compressive stresses in the top layer [2, 32] and obtain a top layer free of abrasive contaminants. These features favorably affect a number of functional properties, including the tribological wear [5, 24], fatigue strength [33, 34] and corrosion resistance [28, 37].

In [6] the authors conducted tests of burnishing 6061 series aluminium alloy. They proved that by using appropriate machining parameters, it is possible to reduce the friction coefficient of friction couples by 48% and weight loss by 60-80%.. Hamadache et al. [10] applied the burnishing process to Rb40 steel. The tests have demonstrated the severalfold increase in wear resistance, compared to samples after machining processes. The results of the research carried out by the authors of [18] demonstrate the beneficial effect of burnishing on the reduction of the friction coefficient and the wear of polymers in relation to the non-burnished samples. Polyurethane and polyformaldehyde were tested. The reduction of the friction coefficient in tribological tests for both polymers reached a maximum of 32%, while the wear was reduced by a maximum of 38%, when compared to samples subjected to machining processes. The degree of reduction of both the friction coefficient and the wear depended strongly on the initial value of the surface roughness. Revankar et al. [26] studied the tribological properties of the Ti-6Al-4V titanium alloy. They tested surfaces following the machining and burnishing processes, for a range of initial process input parameters. It was proved that for the most favorable parameters there was a 52% reduction in wear and a 64% reduction in friction coefficient compared to machined samples. Whereas the authors of [24] examined the wear resistance of two-phase HSLA steel alloys subjected to burnishing process. It was observed that a greater reduction in wear was recorded in the case of samples burnished with a greater load.. The purpose of the work [35] was to determine the possibilities of limiting 316L steel damage caused by impact wear without disturbing the material structure. Burnishing was selected as the finishing treatment of the steel surface. Studies have demonstrated a 53-62% reduction in wear traces compared to the non-burnished samples. Janczewski et al. [12] studied low density and high molecular weight (LDPE) polyethylene samples. Milling and burnishing processes were selected as the finish. Burnishing proved to modify the LDPE surface and dramatically reduce wear by approx. 58% compared to milled samples.

During the studies on the influence of burnishing on the tribological properties of friction couples, in their vast majority, diamond or diamond composite is used as the burnishing element, which makes it a relatively expensive solution. There is little mention of the use of cheaper burnishing elements applied in particular to numerically controlled machine tools to improve these properties. The possibility of wider application of ceramic materials [3, 26] for their improvement would undoubtedly lead to a reduction of costs associated with the burnishing process, in particular in the case of obtaining surface quality at a level similar to that obtained as a result of using diamond tooling. It would, at the same time, eliminate or reduce labor-intensive finishing operations such as honing, grinding, polishing or superfinishing.

The present work presents the research on the impact of burnishing process carried out with use of various ceramics on selected tribological properties of steel - steel couples. In addition, our comparative analysis included standard machining processes used in machine construction, such as polishing, grinding and lapping.

2. Experimental

Tribological tests were carried out on a T-11 pin/ball - disc tribological tester, in the ball - disc configuration. The tribological couple was a stationary ball bearing with a hardness of 62±2 HRC and a disc made of 42CrMo4 steel with a hardness of 40±2 HRC. The finishing treatment of the discs consisted of burnishing processes completed with various types of ceramic balls (Al_2O_3, SiC, WC). Burnishing was performed using a vertical Hass VF-3 machining center using 3 different pressure forces of the tool: 30, 70 and 100 N. A single tool pass was used at a constant feed rate of 0.05 mm. In addition, discs in which grinding, polishing and lapping processes were the finishing treatment were tested. Isometric views of the surfaces of selected discs are presented in Fig. 1. In turn, Table 1 summarizes selected parameters of the surface topography of the tested discs [20]. For discs after the burnishing process the following designations were used:

- Al_{30}, Al_{70} and Al_{100} when the burnishing element was Al_2O_3 ceramics (indexes denote the pressure force applied during burnishing),
- SiC_{30}, SiC_{70}, SiC_{100} when the burnishing element was SiC ceramics,
- WC_{30}, WC_{70}, WC_{100} when the burnishing element was WC ceramics.

In addition, the grinding, polishing and lapping processes were designated 'SZ', 'POL' and 'DOC' respectively. The sliding distance in all variants was 282.6 m, which corresponded to 30 minutes of the test. The tests were carried out at a load of 9.81 N and at three sliding speeds: 0.16; 0.32 and 0.48 m/s. During the tests, the friction force was measured, while after the tests the amount of wear was determined using a Talysurf CCI Light white light interferometer. Measurements of wear were made at four positions 90^0 apart, obtaining areas of 3.3 mm x 3.3 mm. Then, profiles were generated perpendicular to the wear track and the area of wear was calculated using TalyMap Gold 6.0 software. The next step was to calculate the volumetric wear of discs according to the formula (1):

$$VD = \pi dS\left[\mathrm{m}^3\right] \quad (1)$$

where:
d – diameter of the wear track [mm],
S – cross-sectional area of the wear track [mm^2].

All tests were repeated at least 3 times. Residual stress measurements were carried out with use of a portable Xstress 3000 G3R X-ray diffractometer. The measurements used the sin 2ψ method [7] during which the angle of incidence ψ was in the range of −45° to + 45°, divided into 7 tilt positions. Exposure time was set to 40 s. The X-ray

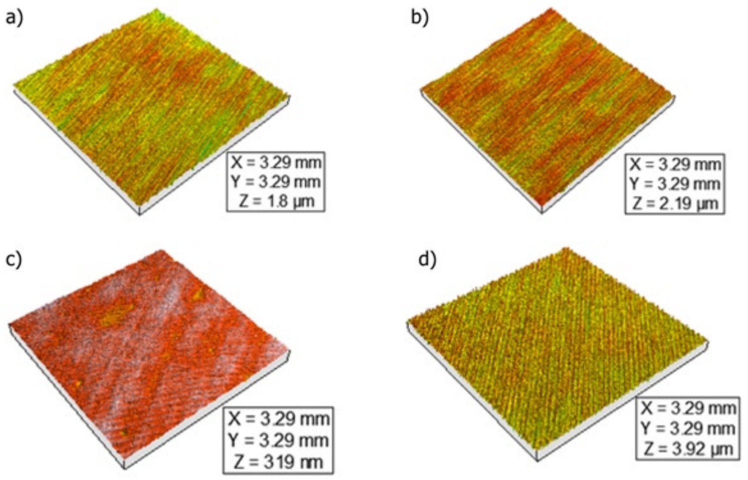


Fig. 1. Isometric views of selected discs: a) burnished sample Al_{70}, b) burnished sample SiC_{70}, c) polished sample, d) ground sample

Table 1. Selected surface topography parameters of tested discs

	Sq	Ssk	Sku	Sz	Sal	Spd	Sk	Spk	Svk
Al_{30}	0.28	-0.898	4.36	2.84	0.0256	750	0.527	0.131	0.461
Al_{70}	0.142	-1.24	6.12	1.8	0.0172	612	0.336	0.1741	0.325
Al_{100}	0.186	-1.116	4.98	2.44	0.0185	672	0.426	0.148	0.389
SiC_{30}	0.227	-1.07	5.14	2.56	0.214	304	0.385	0.1335	0.395
SiC_{70}	0.117	-0.801	4.71	2.19	0.115	493	0.222	0.0538	0.17
SiC_{100}	0.192	-0.569	3.89	1.98	0.166	468	0.351	0.0944	0.301
WC_{30}	0.207	-0.951	5.9	3.76	0.0107	790	0.445	0.144	0.35
WC_{70}	0.3	-0.321	3.86	3.46	0.379	574	0.465	0.119	0.387
WC_{100}	0.226	-0.869	4.11	2.86	0.189	687	0.458	0.123	0.379
SZ	0.258	-0.277	3.81	3.92	0.0086	843	0.603	0.123	0.333
POL	0.0189	-0.17	2.9	0.319	0.214	1170	0.0319	0.01	0.0181
DOC	0.094	-0.59	4.39	1.3	0.0264	1020	0.198	0.0608	0.137

penetration depth was set to 10 μm and the XTronic software was used to compute the residual stresses. For each sample, residual stress was determined in 2 directions perpendicular and parallel to machining traces.

Microhardness measurements were carried out on a Brivisor KL2 microhardness tester with HME measuring electronics by means of static indenter induction using the Vickers method, with a constant load of P = 4.9 N. The impact time of the pyramid indenter with quadrilateral base with a double-wall angle of 136° was ca. 15 s. Surface microhardness of the examined samples was measured on microsections made at an angle of 5°.

3. Results and discussion

Figures 2-4 present the results of the tests. One by one the average values of volumetric wear of the tested disc surfaces for all sliding speeds (Fig. 2), the average values of the friction force obtained after the running-in process (Fig. 3) and the values of the friction force obtained for the sliding distance equal to 30 m (Fig. 4) are presented. Table 2 presents the maximum values of residual stress σ_{max} and microhardness Hm as well as confidence intervals u (σ_{max}), u (Hm) of all samples subjected to tribological tests. Figure 5 shows the cross-sectional areas of wear of selected discs (Al_{70} and POL), and Figure 6 shows examples of friction force curves for these frictional pairs at a sliding speed of 0.16 m/s.

At the lowest sliding speed of v = 0.16 m/s, the largest volumetric wear was observed for the frictional pair with the polished disc. In this case the value of VD parameter was 0.694 mm^3 and corresponded to the highest value of the friction force obtained for the sliding distance equal to 30 m (7.522 N) and the highest average value of the friction force $F_{śr}$ (8.308 N). The smallest value of VD parameter at the sliding speed of v = 0.16 m/s was calculated for the burnished sample Al_{70} and it was 0.439 mm^3. For this sample, the lowest average friction force $F_{śr}$ was also observed, and it was 7.808 N. The friction coefficient μ ranged from 0.79 (Al_{70} sample) to 0.85 (polished sample).

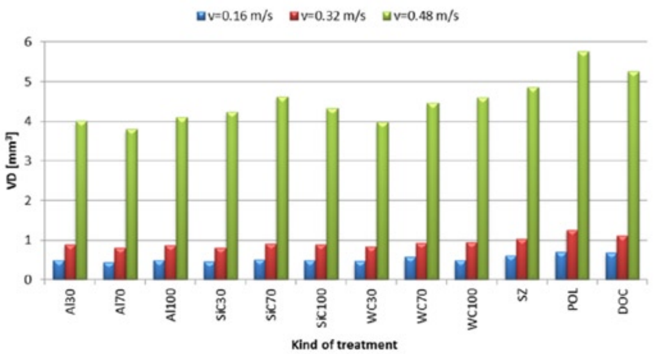


Fig. 2. Volumetric wear of disc sample

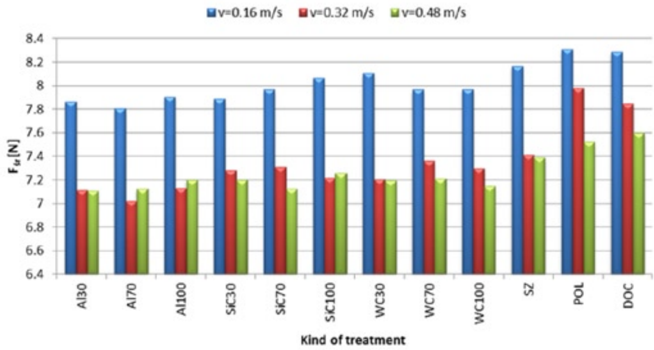


Fig. 3. Average friction force of disc samples

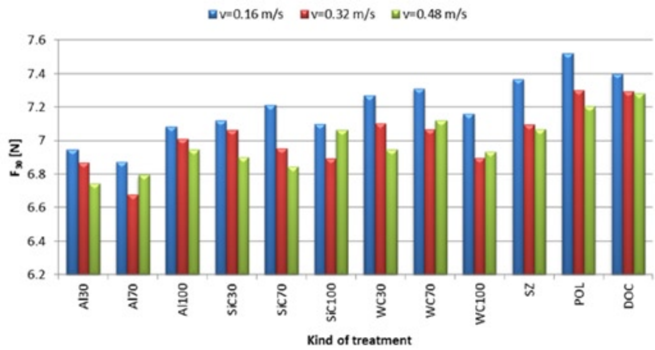


Fig. 4. Friction force of disc samples obtained for the sliding distance equal to 30 m

An increase in the sliding speed to 0.32 m/s caused an increase in the value of volumetric wear of most samples and a slight decrease in the average value of the friction force and the value of the friction force obtained for the sliding distance equal to 30 m. Similar to the lowest sliding speed, the sample Al_{70} had the lowest value of wear volume and it was 0.796 mm^3. In the case of Al_{70} sample the average value of friction force $F_{śr}$ and F_{30} parameter were also the smallest – 7.022 N and 6.677 N respectively. The highest values of measured parameters were obtained for the polished sample. The VD parameter for this sample achieved 1.262 mm^3. The friction coefficient μ de-

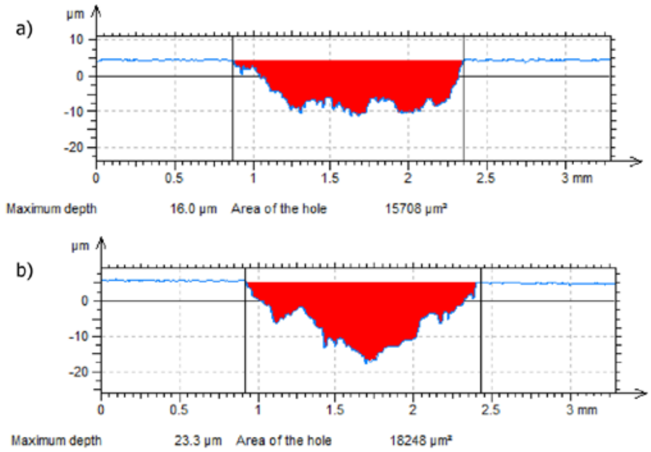


Fig. 5 Cross-sectional areas of worn discs: Al_{70} (a) and POL (b) at the sliding speed of 0.16 m/s,

creased in comparison to the lowest sliding speed and ranged from 0.72 (Al_{70} disc) to 0.81 (polished disc).

An increase in sliding speed to 0.48 m/s caused a very substantial

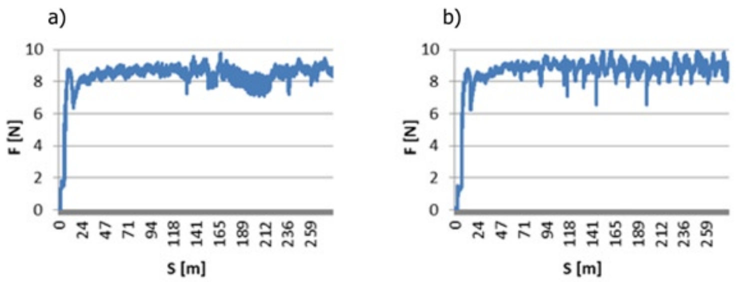


Fig. 6. Friction force versus sliding distance for selected worn discs Al_{70} (a) oraz POL (b) at a sliding speed of 0.16 m/s

increase in the volumetric wear of all samples compared to the wear observed at lower sliding speeds. Al_{70} sample was again characterized by the lowest wear volume at this sliding speed, however the lowest values of $F_{śr}$ and F_{30} were observed for the sample Al_{30}. The VD parameter of the Al_{70} sample was 3.796 mm^3. The highest value of volumetric wear was again measured for the polished sample and it was 5.759 mm^3. In turn, the highest values of the $F_{śr}$ and F_{30} parameters were recorded for the lapped sample and it was 7.598 N and 7.284 N, respectively. The lower value of the friction coefficient was at a similar level as at the sliding speed of 0.32 m/s, while the upper value slightly decreased compared to this speed. The friction coefficient was in the range of 0.72 (Al_{30} sample) to 0.77 (lapped sample).

The highest values of maximum compressive stresses were obtained for the burnished samples SiC_{100} and WC_{100}. The measured values of these discs were –376.1 MPa and –349.6 MPa respectively. In turn, the highest maximum value of surface microhardness was observed for the burnished surface topography Al_{100}. Despite this, mentioned samples were characterized by relatively high volumetric wear of the discs among the samples subjected to the slide burnishing process. The lowest values of the VD parameter were observed for Al_{70} and SiC_{30} discs at the sliding speed of 0.16 m/s, again for Al_{70} and SiC_{30} discs at the sliding speed of 0.32 m/s, as well as for Al_{70} and WC_{30} discs at the sliding speed of 0.48 m/s.

Table 2. Maximum values of residual stress σ_{max} and microhardness Hm of the surface layer

Sample	σ_{max} [MPa]	u (σ_{max}) [MPa]	Hm [HV]	u (Hm) [HV]
Al_{30}	–212.6	±18.6	380.1	±11.6
Al_{70}	–298.1	±17.8	389.4	±14.2
Al_{100}	–309.8	±26.2	404.2	±14.8
SiC_{30}	–270.3	±18.6	399.6	±13.4
SiC_{70}	–301.6	±24.8	388.2	±11.6
SiC_{100}	–376.1	±28.6	399.7	±12.2
WC_{30}	–288.7	±26.1	382.4	±13.9
WC_{70}	–344.6	±28.7	403.5	±12.2
WC_{100}	–349.4	±24.4	398.8	±11.4
SZ	–37.2	±8.7	381.4	±12.2
POL	–179.2	±12.4	385.2	±14.6
DOC	–96.8	±10.2	384.1	±13.8

To illustrate the wear mechanism of worn surfaces, a surface topography analysis and scanning microscope (Vega3) analysis were performed. Surface topography analysis showed the formation of a one-directional texture after tribological tests. The values of Str (texture aspect ratio of the surface) were in the range of 6.98–13.06% and they were characteristic for anisotropic surfaces after abrasive wear. Figure 7 shows isometric views of Al_{70} and POL discs after tribological tests, as well as slices of their worn surfaces. For the Al_{70} sample the Str parameter equaled to 8.44%, and for the polished sample it was 9.13%.

Also, SEM analysis (Figure 8) confirms abrasion as the dominant wear mechanism. As a result, a complete change in the surface topography was observed. The tracks of wear on the disc surfaces include elongated craters in the sliding direction, and smoothed areas with longitudinal grooves resulting from plastic deformation. The presence of delaminations on the worn surface of discs was also observed.

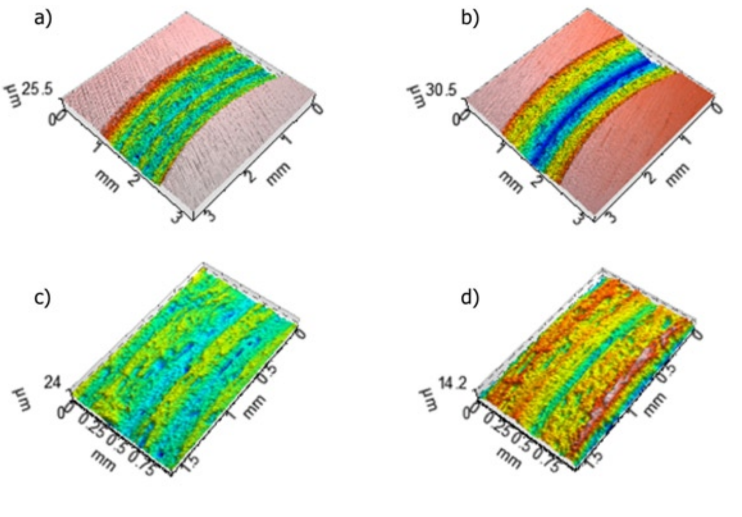


Fig. 7. Isometric views of the discs Al_{70} (a) and POL (b) after tribological tests and details of their worn surfaces Al_{70} (c) i POL (d)

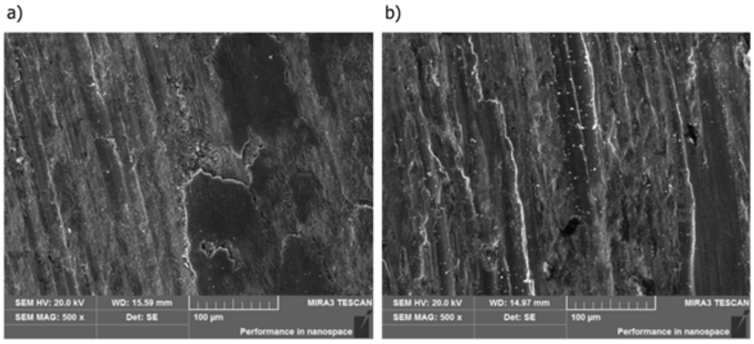


Fig. 8. Worn surfaces of the discs Al_{70} (a) and POL (b) at the sliding speed of v = 0.16 m/s

The coefficient of the linear correlation R was used to search for dependency between tribological parameters and parameters characterizing the surface topography. The value of the correlation coefficient ranges from –1 to 1. The higher its absolute value, the stronger the linear relationship between the variables. The value R = 1 or R = −1 indicates the complete linear dependency between the features, while R = 0 indicates lack of such relation. A strong linear dependency occurring at all sliding speeds was found for VD-Ssk (skewnees) and VD-Sku (kurtosis) pairs. These dependencies are shown in Figure 9. The skewness Ssk, also known as the asymmetry coefficient, characterizes the symmetry of the ordinate distribution of the roughness height relative to the mean plane [15]. According to [27], a negative value of this factor indicates a plateau like surface, while a positive one indicates the predominance of peaks. All tested discs were characterized by a negative value of the Ssk parameter, which ranged from −0.17 (polished disc) to −1.24 (Al_{70} disc). In turn, the Sku parameter is determined by the measure of the peaks or sharpness of the surface height distribution. For the normal distribution of the ordinates Sku = 3 [8]. Samples subjected to tribological tests were characterized by normal ordinate distribution or close to normal one. In the case of volumetric wear of discs and the Ssk parameter, the coefficient of linear correlation R ranged from 0.77 (for v = 0.48 m/s) to 0.83 (for v = 0.16 m/s), and for VD and Sku parameter the coefficient of linear correlation R achieved values between −0.72 (for v = 0.16 m/s) and −0.78 (for v = 0.32 m/s) and it was inversely proportional. The smallest values of volumetric wear of discs were achieved for the minimum values of the Ssk parameter and the maximum values of the Sku parameter. Negative skewness can improve contact conditions by reducing the plasticity index, what can speed up the reduction in wear volume. Surfaces with a low value of Ssk and a high value of Sku can be „traps" for wear particles. In the case of the research in dry friction conditions and at ambient temperatures, wear particles typically have from about 10 to 100 µm. It seems that some of the particles were so small that the valleys of surface

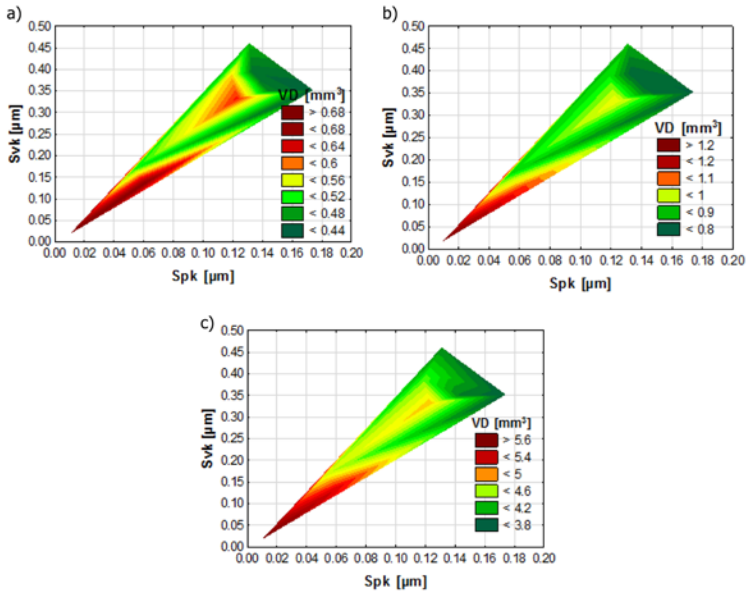


Fig. 10. Dependencies between wear volume VD and Spk i Svk parameters at the sliding speed of: v = 0.16 m/s (a), v = 0.32 m/s (b), v = 0.48 m/s (c)

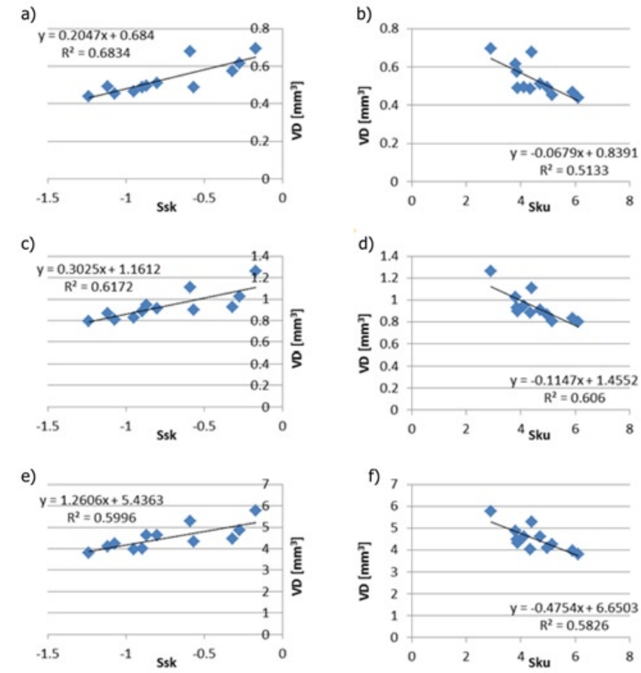


Fig. 9. Dependencies between wear volume VD and Ssk (a, c, e) and Sku (b, d, f) parameters at the sliding speed of: a, b) v = 0.16 m/s; c, d) v = 0.32 m/s; e, f) v = 0.48 m/s

topography acted as traps for some of the wear products and helped to reduce the wear intensity. The situation could be different in the case of tests at elevated temperature, and comprehensive studies on the structure of wear products resulting from the friction process were conducted by the authors [39, 40].

A strong linear dependency occurring at all sliding speeds was also found for volumetric wear and parameters characterizing the Abbott-Firestone curve: Spk (reduced peak height), Svk (reduced valley depth) and Sk (core roughness depth). The relations between Spk and Svk are presented in Fig. 10. The areas marked in red indicate a higher value of the VD parameter, and the green one (in particular dark green) with less wear of the discs. Analyzing the obtained results, it can be seen that the increase in the Spk and Svk parameters corresponded to a decrease in the wear volume of the discs. The Svk parameter allows the evaluation of the lubricating properties of surfaces and is a measure of the ability to maintain fluid through the mating surfaces. In turn, higher values of the Spk parameter characterize surfaces with high peaks, which makes the area of initial contact relatively small and the force applied per unit of surface large. Therefore, the Spk parameter may represent the nominal height of the material that will be removed at the initial stage of the operation - running in [19]. An increase of Spk parameter leads to a reduction in the real contact surface, which may limit the impact of adhesive effects. The authors of [29] suggest that the Spk/Svk ratio is more important than the value of these parameters separately. In all examined surface topographies, the Svk parameter was greater than Spk, which made the tribological couples tend to reduce the coefficient of friction (especially with a negative value of the Ssk parameter). The dependency between the Sk parameter and the wear volume of the discs seem to be particularly important. The Sk parameter controls the tribological properties of the elements after the running-in period. In the case of discs with a small roughness height this is particularly important because the running-in period in all tested configurations did not exceed 5 minutes. Therefore, this parameter can be taken into account when planning the surface topography with the desired tribological properties.

A strong dependency was also observed between volumetric wear and the Spd parameter (peak density). This parameter determines the number of peaks by the unit area. The greater the value of the Spd parameter, the greater the bearing surface [23], especially when the Ssk parameter is negative.

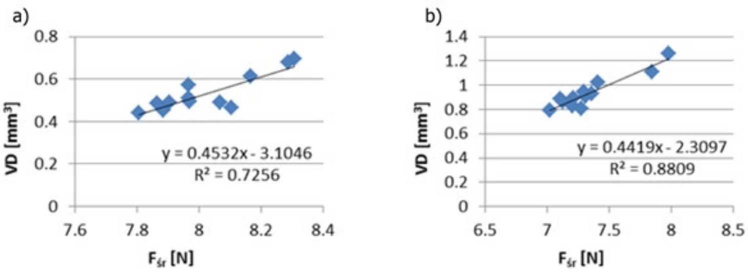


Fig. 11. Dependencies between wear volume VD and average value of friction force $F_{śr}$ at the sliding speed of: v = 0.16 m/s (a) and v = 0.32 m/s (b)

The volumetric wear of the discs and the average value of the friction force were also characterized by a strong linear dependency (Fig. 11). Depending on the sliding speed, the coefficient of linear correlation was: R = 0.85 at v = 0.16 m/s; R = 0.94 at v = 0.32 m/s and R = 0.84 at v = 0.48 m/s. Also the values of friction force obtained for the sliding distance equal to 30 m were strongly correlated with the wear volume of the discs. In this case, depending on the sliding speed, the coefficient of linear correlation between VD and F_{30} achieved: R = 0.85 at v = 0.16 m/s; R = 0.73 at v = 0.32 m/s and R = 0.78 at v = 0.48 m/s. In turn, the wear volume of discs was not significantly correlated with the maximum value of residual stresses (coefficient of linear correlation ranged from −0.54 to −0.64) and microhardness (coefficient of linear correlation in the range: −0.25 ÷ −0.34).

4. Conclusions

Based on the tests carried out, we can conclude that the increase in sliding speed led to an increase in the wear volume of discs subjected to tribological tests. Of all samples, the lowest values of the VD parameter were recorded in the case of Al_{70} oraz SiC_{30} burnished discs. The wear volume of the discs was correlated with the average value of the friction force and also with the value of the friction force obtained for the sliding distance equal to 30 m.

A number of relationships between the parameters characterizing surface topography of the discs and wear volume were also found. A strong linear relationship was observed for the parameters characterizing the Abbot-Firestone curve (Sk, Svk, Spk) as well as for the skewnees and kurtosis of the surfaces (Ssk and Sku).

The tests proved the beneficial effect of the slide burnishing process (regardless of the ceramics applied) on the reduction of volumetric wear in friction pairs compared to other popular finishing treatments used in machine construction.

References

1. Bara M, Skoneczny W, Kaptacz S. A Tribological properties of ceramic-carbon surface layers obtained in electrolytes with a different graphite content. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2008; 4: 66-70.
2. Chomienne V, Valiorgue F, Rech J, Verdu C. Influence of ball burnishing on residual stress profile of a 15-5PH stainless steel. CIRP Journal of Manufacturing Science and Technology 2016; 13: 90-96, https://doi.org/10.1016/j.cirpj.2015.12.003.
3. Dzierwa A, Markopoulos A P. Influence of ball-burnishing process on surface topography parameters and tribological properties of hardened steel. Machines 2019; 7(1): 11, https://doi.org/10.3390/machines7010011.
4. Dzierwa A, Pawlus P, Reizer R. The effect of ceramic tribo-elements on friction and wear of smooth steel surfaces. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology 2019; 233(3): 456-465, https://doi.org/10.1177/1350650118780779.
5. El-Tayeb N S M. Frictional behaviour of burnished copper surfaces under dry contact conditions. Engineering Research Bulletin 1994; HU Cairo: 171-184.
6. El-Tayeb N S M, Low K O, Brevern P V. Enhancement of surface quality and tribological properties using ball burnishing process. Machining Science and Technology 2008; 12: 234-248, https://doi.org/10.1080/10910340802067536.
7. Fitzpatric M E, Fry A T, Holdway P, Kandil F A, Shackleton J, Suominen L L. Determination of Residual Stresses by X-ray Diffraction - Issue 2. A National Measurement Good Practice Guide No. 52: National Physical Laboratory, 2005.
8. Gadelmawla E S, Koura M M, Maksoud T M A, Elewa I M, Soliman H H. Roughness parameters. Journal of Materials Processing Technology 2002; 123: 133-145, https://doi.org/10.1016/S0924-0136(02)00060-2.
9. Ganesh B K C, Sha W, Ramanaiah N, Krishnaiah A. Effect of shot peening on sliding wear and tensile behavior of titanium implant alloys. Materials & Design 2014; 56: 480-486, https://doi.org/10.1016/j.matdes.2013.11.052.
10. Hamadache H, Laouar L, Zeghib N E, Chaoui K. Characteristics of Rb40 steel superficial layer under ball and roller burnishing. Journal of Materials Processing Technology 2006; 180: 130-136, https://doi.org/10.1016/j.jmatprotec.2006.05.013.
11. Hassan A M. The effects of ball and roller burnishing on the surface roughness and hardness of some non-ferrous metals. Journal of Materials Processing Technology 1997; 72: 385-391, https://doi.org/10.1016/S0924-0136(97)00199-4.
12. Janczewski Ł, Toboła D, Brostow W, Czechowski K, Hagg Lobland K E, Kot M, Zagórski K. Effects of ball burnishing on surface properties of low density polyethylene. Tribology International 2016; 93: 36-42, https://doi.org/10.1016/j.triboint.2015.09.006.
13. Jedliński R. Diagnostyka w procesie wyceny wartości pojazdów samochodowych i ustalaniu przyczyn oraz następstw uszkodzeń. Diagnostyka 2005, 33: 65-70.
14. Jerez-Mesa R, Gomez-Gras G, Travieso-Rodriguez J A. Surface roughness assessment after different strategy patterns of ultrasonic ball burnishing. Procedia Manufacturing 2017; 13: 710-717, https://doi.org/10.1016/j.promfg.2017.09.116.

15. Jerez-Mesa R, Travieso-Rodríguez J A, Landon Y, Dessein G, Lluma-Fuentes J, Wagner V. Comprehensive analysis of surface integrity modification of ball-end milledTi-6Al-4V surfaces through vibration-assisted ball burnishing. Journal of Materials Processing Technology 2019; 267: 230-240, https://doi.org/10.1016/j.jmatprotec.2018.12.022.
16. Korzynski M, Dzierwa A, Pacana A, Cwanek J. Fatigue strength of chromium coated elements and possibility of its improvement with ball peening. Surface and Coatings Technology 2009; 204: 615-620, https://doi.org/10.1016/j.surfcoat.2009.08.049.
17. Kovács Z F, Viharos Z J, Kodácsy J. Determination of the working gap and optimal machining parameters for magnetic assisted ball burnishing. Measurement 2018; 118: 172-180, https://doi.org/10.1016/j.measurement.2018.01.033.
18. Low K O, Wong K J. Influence of ball burnishing on surface quality and tribological characteristics of polymers under dry sliding conditions. Tribology International 2011; 44: 144-153. https://doi.org/10.1016/j.triboint.2010.10.005
19. Niemczewska-Wojcik M. Wear mechanisms and surface topography of artificial hip joint components at the subsequent stages of tribological tests. Measurement 2017; 107: 89-98, https://doi.org/10.1016/j.measurement.2017.04.045.
20. PN-EN ISO 25178-2:2012 Geometrical Product Specifications (GPS) - Surface texture: Areal - Part 2: Terms, definitions and surface texture parameters.
21. PN-EN ISO 4516:2004 Powłoki metalowe i inne nieorganiczne. Badania mikrotwardości metodą Vickersa i Knoopa.
22. Priyadarsini Ch, Venkata Ramana V S N, Aruna Prabha K, Swethaa S. A review on ball, roller, low plasticity burnishing process. Materials Today: Proceedings 2019; 18(7): 5087-5099, https://doi.org/10.1016/j.matpr.2019.07.505.
23. Pu Y, Zhao Y, Zhang H, Zhao G, Meng J, Song P. Study on the three-dimensional topography of the machined surface in laser-assisted machining of Si3N4 ceramics under different material removal modes. Ceramics International 2019; In press, Available online 14 November 2019, https://doi.org/10.1016/j.ceramint.2019.11.017.
24. Rao D S, Suresh H H, Komaraiah M, Kempaiah U N. Investigations on the effect of ball burnishing parameters on surface hardness and wear resistance of HSLA dual-phase steels. Materials and Manufacturing Processes 2008; 23: 295-302, https://doi.org/10.1080/10426910801937306.
25. Revankar G D, Shetty R, Rao S S, Gaitonde V N. Analysis of surface roughness and hardness in ball burnishing of titanium alloy. Measurement 2014; 58: 256-268, https://doi.org/10.1016/j.measurement.2014.08.043.
26. Revankar G D, Shetty R, Rao S S, Gaitonde V N. Wear resistance enhancement of titanium alloy (Ti-6Al-4V) by ball burnishing process. Journal of Materials Research and Technology 2017; 6: 13-32, https://doi.org/10.1016/j.jmrt.2016.03.007.
27. Sadowski Ł, Czarnecki S, Hoła J. Evaluation of the height 3D roughness parameters of concrete substrate and the adhesion to epoxy resin. International Journal of Adhesion and Adhesives 2016; 67: 3-13, https://doi.org/10.1016/j.ijadhadh.2015.12.019.
28. Saldaña-Robles A, Plascencia-Mora H, Aguilera-Gómez E, Saldaña-Robles A, Marquez-Herrera A, Diosdado-De la Peña J A. Influence of ball-burnishing on roughness, hardness and corrosion resistance of AISI 1045 steel. Surface and Coatings Technology 2018; 339: 191-198, https://doi.org/10.1016/j.surfcoat.2018.02.013.
29. Sedlacek M, Podgornik B, Vizintin J. Influence of surface preparation on roughness parameters, friction and wear. Wear 2009; 266: 482-487, https://doi.org/10.1016/j.wear.2008.04.017.
30. Swirad S, Wdowik R. Determining the effect of ball burnishing parameters on surface roughness using the Taguchi method. Procedia Manufacturing 2019; 34: 287-292, https://doi.org/10.1016/j.promfg.2019.06.152.
31. Swirad S, Wydrzynski D, Nieslony P, Krolczyk G M. Influence of hydrostatic burnishing strategy on the surface topography of martensitic steel. Measurement 2019; 138: 590-601, https://doi.org/10.1016/j.measurement.2019.02.081.
32. Teimouri R, Amini S, Bagheri Bami A. Evaluation of optimized surface properties and residual stress in ultrasonic assisted ball burnishing of AA6061-T6. Measurement 2018; 116: 129-139, https://doi.org/10.1016/j.measurement.2017.11.001.
33. Travieso-Rodríguez J A, Jerez-Mesa R, Gómez-Gras G, Llumà-Fuentes J, Casadesús-Farràs O, Madueño-Guerrero M. Hardening effect and fatigue behavior enhancement through ball burnishing on AISI 1038. Journal of Materials Research and Technology 2019, 8(6): 5639-5646, https://doi.org/10.1016/j.jmrt.2019.09.032.
34. Valiorgue F, Zmelty V, Dumas M, Chomienne V, Verdu C, Lefebvre F, Rech J. Influence of residual stress profile and surface microstructure on fatigue life of a 15-5PH, Procedia Engineering 2018; 213: 623-629, https://doi.org/10.1016/j.proeng.2018.02.058.
35. Yilmaz H, Sadeler R. Impact wear behavior of ball burnished 316L stainless steel. Surface and Coatings Technology 2019; 363: 369-378, https://doi.org/10.1016/j.surfcoat.2019.02.022.
36. Zaleski K. The effect of vibratory and rotational shot peening and wear on fatigue life of steel. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19(1): 102-107, https://doi.org/10.17531/ein.2017.1.14.
37. Zhang P, Liu Z. Enhancing surface integrity and corrosion resistance of laser cladded Cr-Ni alloys by hard turning and low plasticity burnishing. Applied Surface Science 2017; 409: 169-178, https://doi.org/10.1016/j.apsusc.2017.03.028.
38. Żurowski W. Structural factors contributing to increased wear resistance of steel friction couples. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2012; 14(1): 19-23.
39. Żurowski W, Brzózka K, Górka B. Analysis of surface layers and wear products by Mössbauer spectral analysis. Wear 2013; 297(1): 958-965, https://doi.org/10.1016/j.wear.2012.10.012.
40. Żurowski W, Brzózka K, Górka B. Structure of friction products and the surface of tribological system elements. Nukleonika 2013; 58: 99-103.

Andrzej DZIERWA
Lidia GAŁDA
Rzeszow University of Technology
Faculty of Mechanical Engineering and Aeronautics
ul. Powstańców Warszawy 8, 35-959 Rzeszow, Poland

Mirosław TUPAJ
Rzeszow University of Technology
Faculty of Mechanics and Technology
ul. Kwiatkowskiego 4, 37-450 Stalowa Wola, Poland

Kazimiera DUDEK
University of Rzeszow
Centre for Innovative Technologies,
ul. Pigonia 1, 35-310 Rzeszow, Poland

E-mails: adzierwa@prz.edu.pl, lgktmiop@prz.edu.pl,
mirek@prz.edu.pl, kaziadudek@o2.pl

Article citation info:
RODRIGUES J, COSTA I, TORRES FARINHA J, MENDES M, MARGALHO L. Predicting motor oil condition using artificial neural networks and principal component analysis. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 440–448, http://dx.doi.org/10.17531/ein.2020.3.6.

João RODRIGUES
Inês COSTA
J. TORRES FARINHA
Mateus MENDES
Luís MARGALHO

PREDICTING MOTOR OIL CONDITION USING ARTIFICIAL NEURAL NETWORKS AND PRINCIPAL COMPONENT ANALYSIS

PROGNOZOWANIE STANU OLEJU SILNIKOWEGO ZA POMOCĄ SZTUCZNYCH SIECI NEURONOWYCH I ANALIZY SKŁADOWYCH GŁÓWNYCH

The safety and performance of engines such as Diesel, gas or even wind turbines depends on the quality and condition of the lubricant oil. Assessment of engine oil condition is done based on more than twenty variables that have, individually, variations that depend on the engines' behaviour, type and other factors. The present paper describes a model to automatically classify the oil condition, using Artificial Neural Networks and Principal Component Analysis. The study was done using data obtained from two passenger bus companies in a country of Southern Europe. The results show the importance of each variable monitored for determining the ideal time to change oil. In many cases, it may be possible to enlarge intervals between maintenance interventions, while in other cases the oil passed the ideal change point.

Keywords: *condition monitoring, oil analysis, multivariate analysis, predictive maintenance.*

Bezpieczeństwo i wydajność silników takich, jak silniki Diesla czy gazowe, a nawet turbiny wiatrowe, zależą od jakości i stanu oleju smarowego. Stanu oleju silnikowego ocenia się na podstawie ponad dwudziestu zmiennych, z których każda ulega wahaniom w zależności od typu i zachowania silnika oraz innych czynników. W niniejszym artykule opisano model, który pozwala na automatyczną klasyfikację stanu oleju, z wykorzystaniem sztucznych sieci neuronowych i analizy składowych głównych. Badania przeprowadzono na podstawie danych uzyskanych od dwóch przewoźników pasażerskich działających na terenie jednego z krajów położonych na południu Europy. Wyniki pokazują, że każda z monitorowanych zmiennych ma znaczenie dla określenia idealnego czasu na wymianę oleju. Podczas gdy w wielu przypadkach w badanych przedsiębiorstwach możliwe było zwiększenie odstępów czasowych między działaniami konserwacyjnymi, w innych, idealny moment wymiany oleju został przekroczony.

Słowa kluczowe: *monitorowanie stanu, analiza oleju, analiza wielowymiarowa, konserwacja predykcyjna.*

1. Introduction

Condition monitoring of engines' oil is a strategic area in the maintenance management field. Replacing the oil too early represents unnecessary unavailability, as well a financial and environmental costs which could be spared. Replacing it too late can impair the oil's ability to protect the engine, therefore increasing the chances of damage and premature ageing of the engine, or even the risk of causing accidents which can endanger people, equipments or vehicles in urban environments. The use of modern tools from data mining and Artificial Intelligence (AI) can contribute to help make the right decision at the right time, thus protecting the environment, the companies' profits and the safety of people and property.

The present paper discusses a methodology to create models to facilitate the process of oil analysis, tested with a dataset for oil of Diesel engines, from urban passenger buses. Preliminary work was already done [14], using Artificial Neural Networks (ANN). In the present research, the neural models were improved and the results are compared with analysis using multivariate systems, namely Principal Component Analysis (PCA). PCA showed the relevance of each variable is different, and some of the variables may even have a negative impact on the predictive power of the ANN.

Data used for the experiments come from two passenger bus companies. Each company provided a dataset, containing results of laboratory analysis of the oils and their classification, according to human experts of a specialized oil analysis company. Data were mined and neural models were created, for both datasets separated and combined.

The remainder of the paper is organized as follows. Section 2 presents a summary of the state of the art. Section 3 describes the datasets used. Section 4 describes the neural networks. Section 5 describes the analysis performed using multivariate systems. Section 7 presents a comparative and critical analysis of the results obtained. Section 8 highlights the main contributions of the present research. Section 9 presents some conclusions and outlines future work.

2. Literature Review

2.1. Condition monitoring of Diesel engines' oil

Condition monitoring of Diesel engines' oil has been subject to study using different approaches, including machine learning methods. Raposo et al. present a study about condition monitoring based on oil in the Diesel engines of a fleet of urban buses. The study shows

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

the evolution of oil degradation and develops a predictive maintenance policy for oil replacement [13]. The methodology presented by the authors considers only some variables of the oils, showing very interesting results about the P-F curve accompaniment. The P-F curve is the interval between the detection of a Potential failure and the actual Failure— that is, the interval where a maintenance team should intervene to prevent a potential failure from happening.

Gajewski & Valis present a study that focuses on heavy transport systems. The types of oils were obtained from several dozens engines of heavy crawlers. The study uses these data with neural networks, in order to identify the patterns that model the system deterioration [5]. Hongxiang et al. [7] use a feed forward neural network to classify different types of oil and their running/not running condition. Parlak et al. [12] use an ANN to predict specific fuel consumption and Diesel engine temperature.

2.2. Online condition monitoring of engine oil

Monitoring the condition or the engine's oil in real time is also a long sought goal. Oil degradation depends on the working time, kilometers driven, the driving speed and habits, type of motor, age of the motor and many other variables which may cause faster or slower degradation. Therefore, for replacement of the oil at the best time, it is necessary to make analysis to determine the state of degradation of the oil. The main problems with laboratory analysis is that they are a laborious process, which requires human intervention. Even though it is only necessary each several km, the idea is to automate the process, thus lowering the costs and reducing the probability of human error.

Accurate online monitoring has two main advantages: i) it reduces downtime necessary to inspect the oil; and ii) it increases chances that the oil will be changed at the best time, not too early and not too late. On the downside, it requires adequate sensors that can be put in the engines, in contact with the oil. The sensors must be robust enough to endure the operating conditions without failing and they must be precise enough to give accurate readings. J. Zhu et al. present a fair review of state of the art sensors and methods for online condition monitoring of engines' oil [18]. The authors classify the sensors in four different groups: electromagnetic, physical, chemical and optical. Electromagnetic sensors measure the dielectric constant of the oil. A second type measures the oil's conductivity. A third type measures magnetic susceptibility. A fourth type measures oil viscosity. As for physical methods, Zhu et al. mention the viscometer, ultra sound, thermal conductivity sensor and ferrography. As for chemical methods, the techniques reviewed include pH measurement and thin-film contaminant monitor. As for optical techniques, they are reflectometry and infrared absorption.

S. Kumar et al. propose a method for condition monitoring oil engine online, based on an optical sensor that transduces oil darkness into electrical resistance [8]. The colour change of the oil is one of the variables that directly correlates to the quality of the oil. Therefore, the authors argue that monitoring the oil colour it is possible to determine the degradation of the oil and change it at the most appropriate time. The method is in part similar to [17], where Yonghui et al., who combine the use of a fibre optic transducer and an inductive sensor. The inductive sensor detects large ferrous and some non-ferrous wear debris, while the optical sensor detects small particles contamining the oil.

El-Hag et al. use features extracted from acousting and radio frequency partial discharge signals to monitor oil condition in power transformers [4]. Pulse width, rise time and frequency components are used to train a neural network to assess the level of degradation of the oil. J. Zhu et al. propose a method for condition monitoring wind turbine oil using commercially available sensors to measure oil viscosity and dielectric constant sensors [19]. The sensor readings are calibrated based on the relationship between particle concentration and oil degradation. X. Zhu et al. propose a method to condition monitor oil using a sensor that detects wear debris by measuring the inductance change of two planar coils wound around a pair of ferrite cores [20]. The method is in part similar to Du et al.'s approach, which also uses inductive sensors to measure metallic, ferrous and non-ferrous, particles in lubricant oil [3].

2.3. Use of PCA and Artificial Neural Networks in condition monitoring

Principal Component Analysis is one method of multivariate analysis very popular for data mining. It is a statistical procedure to transform data, extract features and determine the most important variables of a dataset. Through PCA analysis, it is therefore possible to predict which variables deserve to be monitored and which variables are candidate to be removed without loosing predictive power. Westerholm and Li use PCA to determine the relationship between fuel parameters and the amount of particles in Diesel motor emissions [16]. Capone et al. use PCA to determine the amount of unburned fuel in lubricating oil [1].

Different Artificial Neural Network architectures have also been used for learning and predicting oil condition. Shaban et al. use a cascade of artificial neural networks to predict transformer oil parameters [15]. Niu et al. compare the performance of ANN and Support Vector Machines for predicting motor emissions [11]. Ghobadian et al. use an ANN to model the performance of a diesel engine using waste oil [6]. Li X et al. and Li Y et al. use convolution neural networks to detect gear faults based on different signals, namely sounds produced by the gears [9] and operational parameters [10].

3. Datasets Used

The present research was performed using two datasets, obtained from two different public transportation bus companies, named A and B for that purpose. The datasets contain the results of 21 parameters of the laboratory analysis of the oil, taken from the buses at different stages. Each sample also contains the bus and the oil mileage. That is a total of 23 parameters which are used as input variables for the present analysis. The 23 parameters are: mileage of the bus, mileage of the oil, amount of antifreeze found in the oil, percentage of fuel, Finacheck water content, sooth, nitration, oxidation, sulfation, TBN, viscosity at 100°C, Al, Cr, Cu, Fe, Mo, Na, Ni, Pb, Si, Sn, V and PQ. The variables were normalized and used as inputs to the neural networks and PCA analysis. The datasets also contain the decision of the specialized company, marked as 1, 2 or 3. Decision 1 means that the company decided the oil is in good condition and can be maintained for normal bus operation. Decision 2 means the oil is reaching the point where it needs to be replaced. Decision 3 means the oil has passed the point when it should have been replaced and the bus must be immediately stopped for safety reasons. Dataset A contains a total of 47 samples, obtained from a number of different buses of company A. Dataset B contains a total of 88 samples, obtained from twenty two different buses, four samples from each bus. For the present study it was not possible to obtain larger datasets—a limitation which could not be overcome. Nonetheless, the results obtained for the neural models and PCA were consistent. They were repeated a number of times and they are repeatable in similar circumstances. Many neural models showed good performance and small error in the train and test sets. The ones preferred for analysis were those with better performance in the test set, thus showing the model is general. That shows the results are valid and the method could be scaled up to larger datasets. Since PCA is a factor-analysis method, the adequacy of the datasets for PCA was tested using Kaiser-Meyer-Olkin (KMO) test [2]. KMO test gives a score between 0 and 1, where in general the higher scores mean the dataset contains enough diversity of samples to apply factor analysis. A low score, on the other hand, means there are high correlations between the variables and the results of the factoring process are unreliable. The KMO test gives a score of 0.35 for dataset A, which is very low, and a score of 0.636 for dataset B, which is acceptable, meaning there is more confidence in the factor analysis results for dataset B.

4. Neural Models

In the present research the neural models used were shallow Feed Forward Neural Networks, with one hidden layer of variable width (number of neurons), and one output layer. The hidden neurons used a sigmoid transfer function, which can be a universal approximation, maintaining the output in the range [0, 1]. The output neuron used a linear transfer function (relu), to allow for a wider amplitude of the output and facilitate the learning process. The models were created and tested in MatlabTM1. Training was performed using the Levenberg-Marquardt method. The Mean Squared Error (MSE) and correlation factor R were used for performance assessment. Training was performed with 70 % of the samples, validation with 15 % and test with the remainder 15 %. The results of the training process are variable for each experiment. That happens when the initial weights and bias of the neurons are not set for a specific value, or if the samples for the training and validation sets are chosen randomly at each experiment. Therefore, the results of the experiments presented below are selected from a number of runs. During training, the training process was stopped when the error increased in the validation set for two consecutive epochs, and the best model was retained. The output obtained from the neural models is a floating point number. That is desirable, so that it is interpreted as a measure of the quality of the oil: the lowest the value, the better the quality of the oil. On the other hand, it is also important to map the output to 1, 2 or 3, in order to obtain a model that can be compared to the classification of the human experts, as described in Section 3. So it was mapped in the discrete interval [1, 3] using the following rules: Anything below 1.50 was mapped to 1; Numbers in the interval [1.50, 2.50[were mapped to 2 and everything greater or equal to 2.50 was mapped to 3.

Table 1. R and MSE obtained for different network sizes, with dataset A

Hidden layer size	R (train set)	R (validation set)	R (test set)	R (all dataset)	MSE (all dataset)
1	0.98	0.83	0.78	0.91	0.485
2	0.98	0.76	0.80	0.93	0.234
3	0.99	0.96	0.96	0.98	0.051
4	0.99	0.87	0.88	0.95	0.258
5	0.98	0.91	0.70	0.92	0.229
6	0.95	0.90	0.89	0.92	0.216
7	0.98	0.83	0.98	0.96	0.178
8	0.93	0.83	0.88	0.89	0.291
9	0.92	0.88	0.98	0.91	0.178
10	0.90	0.99	0.91	0.89	0.015

Table 2. Confusion matrices of the errors of the model trained with data from Company A. The model shows two prediction errors for Company A, but 39 errors for Company B

Predicted	Company A			Company B		
3	0	0	**14**	0	0	**4**
2	0	**11**	1	0	**0**	2
1	**20**	1	0	**45**	15	22
Actual	1	2	3	1	2	3

4.1. Model for Company A

In order to get the best possible results, it is important to determine the optimal size of the neural network, so that it is able to abstract and retain as much information as possible without overfitting the training data.

Table 1 shows the R and MSE obtained for a number of neurons between 1 and 10, for dataset A. Many models show good R and MSE. The model with three neurons is one of the best, since it shows a good R for all the dataset and a small MSE. It also shows a good R for the test set, which means it is a good general model, performing well even for data that it has never seen. The number of neurons is small for the number of inputs, but it is probably in line with the size of the dataset, which is also small. The model was trained in three epochs, after which it started to show signs of overfitting and, therefore, further training was rejected.

Table 2 shows the confusion matrices with a summary of the distribution of the errors of the model described above, when applied to both datasets. As the table shows, the predictions of the model for Company A are very close to the desired output. There are only two errors, when the model predicted 2 and the decision was 1 and the model predicted 3 while the company decided 2. In both cases the company was more defensive than the model. It should be mentioned that the decisions made by the company are also prone to human error, so the errors shown are not necessarily problems of the model—they can be because of outliers in the dataset.

When the same model was applied to the data obtained from Company B, there were a total of 39 errors in 88 samples. All the errors apparently happen because Company B was more defensive than the model, showing a very clear trend: the experts rated the oil worse than the model which performed very well for Company A. This proves that Company B replaces the oil, on average, before Company A. That may happen because of different maintenance policies, different motors or different oil brands.

Table 3. R and MSE obtained for different network sizes, with dataset B

Hidden layer size	R (train set)	R (validation set)	R (test set)	R (all dataset)	MSE (all dataset)
1	0.98	0.97	0.86	0.95	0.060
2	0.95	0.93	0.88	0.94	0.110
3	0.99	0.90	0.88	0.96	0.222
4	0.95	0.95	0.83	0.92	0.078
5	0.97	0.86	0.80	0.93	0.161
6	0.86	0.84	0.86	0.86	0.249
7	0.91	0.91	0.91	0.90	0.256
8	0.97	0.89	0.93	0.94	0.144
9	**0.95**	**0.91**	**0.94**	**0.95**	**0.073**
10	0.92	0.89	0.92	0.90	0.184

Table 4. Confusion matrices of the errors of the model trained with data from Company B. The model shows six prediction errors for Company B, but thirty three errors for Company A

Predicted	Company A			Company B		
3	16	12	**14**	0	0	**26**
2	4	**0**	0	4	**15**	2
1	**0**	0	1	**41**	0	0
Actual	1	2	3	1	2	3

4.2. Model for Company B

Table 3 shows the R and MSE obtained from a number of neurons between 1 and 10, for models trained with dataset B. The network with one neuron in the hidden layer shows the lowest MSE. However, it also shows a poor R in the test set, compared to the train and validation sets, meaning the model is a bit overfitted. The model with nine neurons in the hidden layer is better, considering that it shows a high R for all dataset, a high R for the test set and a small MSE. That shows the model is more general for the specific problem being addressed. The best performance for the test and validation sets was obtained at epoch four, and after that the model starts to overfit the data.

Table 4 shows the confusion matrices for the model, applied to datasets B and A. On dataset B, there are six errors of the model, compared to the decision of the company. In four situations the model predicted 2, while the company decided 1. So, the model was more defensive, proposing the bus to replace the oil, while the company decided it was good to circulate. In two cases, the model predicted 1 and the company decided 2.

When the same model was applied to data from Company A, there was a total of thirty three errors. In one situation the company was more defensive: one time the model predicted 1 and the company, apparently, decided to stop the bus. In thirty two situations the model was more defensive: sixteen times the model predicted the bus should be taken out of circulation and the company decided the oil was good, twelve times the model predicted the bus should be taken out of circulation and the company just decided to replace the oil, and four times the model predicted the oil should be replaced and the company decided it was good. The results of this confusion matrix are according to the one shown in Table 2: the companies follow different policies, and Company B replaces the oil much earlier than Company A.

5. Principal Component Analysis

5.1. Introduction

PCA is a statistical procedure used to map a set of correlated variables into a new set of uncorrelated variables, called principal components. The principal components are calculated by decreasing order of importance. The first component is the most important, the last is the less important explanatory variable. Each principal component identified is a linear combination of all the original variables. PCA was applied to the datasets presented above, in order to understand the companies' policies, the state of the oils when the samples were collected for chemical analysis, as well as to determine which variables are more important to measure for correct assessment of the situation of the oil. The PCA experiments and analysis were performed using R Studio software.

5.2. Loadings for each variable of the dataset

Table 5 shows results of the PCA analysis for datasets A and B. As the table shows, Si, Fe, Al and Cr contents are the four most important variables, with loadings above 0.7, for dataset A.

For dataset B, Fe, Soot and Cr are the top three variables, and the only ones with score above 0.7. Five of the top ten variables are related to oil status and five are related to wear and contamination.

5.3. Company A

For Company A, the most important components are polluting metallic agents, which are generated by motor wear: slip wear, wear due to friction, wear due to metal fatigue, and wear due to cutting. This means the bus motors suffer a lot of wearing, being advisable the use of additives to reduce friction, oil leaks and even fumes. The use of the right additives might also increase motor expected life.

Table 6 shows the calculated percentage of deterioration of the oil, for each oil sample, as well as the average deterioration. The percentage of deterioration was obtained multiplying the PCA loadings by the normalized value of each variable. So a kind of weighted average is obtained and then compared to the reference values obtained from the oil datasheet.

As the table shows, four different oils are already well beyond their expected useful life, based on manufacturer recommendations. More than 41 % of the samples have passed 60 % degradation. Another relevant aspect is that there is a large variability between samples. The standard deviation in the dataset is 0.312. In general, Company A apparently has a poor maintenance policy. On one hand, the motors suffer a lot of wearing when the oils are still used beyond their reference limits. That can damage the motors and reduce their useful life. On the other hand, some oils may be changed while they are still good, causing unnecessary financial and environmental costs. In two

Table 5. Most important variables, according to PCA analysis, for datasets A and B. The ten first variables that are in both datasets are highlighted in bold

Order of Relevance	Variable	Loading	Variable	Loading
1	**Si Content**	0.872	**Fe Content**	0.889
2	**Fe Content**	0.864	**Soot**	0.835
3	**Al Content**	0.789	**Cr Content**	0.781
4	**Cr Content**	0.729	Viscosity at 100°C	0.689
5	**Sn Content**	0.668	**Sn Content**	0.682
6	**PQ Index**	0.551	**Cu Content**	0.611
7	Ni Content	0.425	Pb Content	0.571
8	**Soot**	0.441	**Sulfation**	0.507
9	**Oxidation**	0.412	Nitration	0.496
10	V Content	0.376	**Oxidation**	0.488
11	**Cu Content**	0.282	**Al Content**	0.482
12	**Sulfation**	0.266	**Si Content**	0.423
13	Mo Content	0.166	**PQ Index**	0.32
14	Pb Content	0.134	Na content	0.166
15	Fuel	0.132	Antifreeze	0.162
16	Na Content	0.118	Water Content	0.127
17	Viscosity	0.089	Mo Content	0.072
18	TBN	0.069	V Content	0.02
19	Nitration	-0.003	Ni Content	-0.008
20	Water content	-0.14	Fuel Content	-0.134
21	Antifreeze	-0.142	TBN Content	-0.395

cases the oils are apparently at just 15 % of their useful life when they were replaced, according to the results of the laboratorial analysis.

5.4. Company B

Table 6 shows the percentage of degradation of the oil samples. As the table shows, none of the samples is out of the limits proposed by the manufacturer, showing the oil is in good condition to protect the Diesel engines. In general the dataset is also more homogenous, with most of the variables near the average value. The standard deviation is 0.126, which is much lower than the standard deviation calculated for dataset A.

Table 8 shows tha average deterioration of oil for each of the twenty two buses. The average is always in the range 30 – 59 %. Bus 1814 shows the lowest average of all. Bus 2160 shows the highest average, with 58.65 %, which is already a very high value. In general, however, the buses of company B show a low level of wear, with the five variables related to wearing and contamination within the ten highest principal components because of their importance in oil deterioration.

6. Neural model with reduced dimensionality

6.1. Merging datasets

Since both datasets contain the same 23 input variables and three-levels output variable, they are fit to be merged together, aiming to produce one more general model. However, since data come from different sources, it is important to avoid skewing the results towards the policies of one of the companies, for the previous analysis have shown that the companies follow different policies and the datasets may have outliers. The problem can be seen as the typical problem of imbalanced datasets, which is solved using techniques such as oversampling the less frequent data or undersampling the most frequent data. In the present case, considering the small dimension of the dataset, data from Company A were oversampled choosing randomly additional samples from dataset A after cross validation.

The merged dataset was named AB and it contains a total of 176 samples, 88 from each dataset. The neural models used for experiments with dataset AB all contained seven neurons in the hidden layer. The number of inputs, however, varied: i) one of the experiments was run with all the 21 laboratorial variables and the mileage as inputs for the neural network; and ii) a second experiment was run using as inputs to the neural network just the ten variables highlighted in Table 5.

Table 9 shows the R and MSE obtained for different models. As the table shows, the model with 23 input variables was very good, with a very small MSE. The model with just 12 inputs is also very good, with R 0.94 for all the dataset and 0.84 for the test set. The table shows that reducing the number of variables it was still possible to generate a good neural model, with good R and a small MSE of just 0.15

Table 10 shows the confusion matrices of the errors counted when simulating the two models trained with dataset AB. The model trained with all the 23 inputs generates 13 prediction errors: five for samples of company B and 8 for company A. .

The model trained with just 12 input variables produces a total of 14 prediction errors: 7 for each company. Those results also show that the model trained with just 10+2 variables seems more general than the model trained with all the variables, for it generates the same number of prediction errors for each company. The smaller model, retaining less information, still shows good performance marks and is perhaps the most general, as discussed in Section 7.

Table 6. Percentage of deterioration of the oil for dataset A

Sample	Bus #	% Deterioration	Sample	Bus #	% Deterioration
1	122	31.7	25	267	66.7
2	122	13.6	26	270	32.5
3	203	49.9	27	270	54.2
4	203	18.8	28	270	89.1
5	214	52.1	29	282	54.6
6	214	150.6	30	282	51.7
7	214	35.6	31	283	69.0
8	214	17.6	32	283	70.7
9	219	13.4	33	289	73.1
10	246	79.1	34	289	85.3
11	246	64.0	35	290	49.3
12	247	83.6	36	294	55.8
13	248	73.6	37	294	54.3
14	249	46.5	38	297	62.3
15	251	54.1	39	209	72.5
16	252	69.2	40	209	73.3
17	254	162.1	41	301	31.0
18	259	128.0	42	301	86.8
19	260	88.3	43	301	31.7
20	265	118.3	44	304	66.4
21	266	48.3	45	304	51.6
22	266	45.6	46	304	54.8
23	266	63.7	47	304	35.7
24	267	55.2	**Average**		62.5

Table 7. Percentage of deterioration (Det.) of the oil for dataset B

Sample	Bus #	% Det.	Sample	Bus #	% Det.	Sample	Bus #	% Det.
1	2175	31.1	31	1737	29.1	61	2127	64.6
2	2175	30.4	32	1737	28.4	62	2127	56.7
3	2175	38.7	33	2148	84.2	63	2127	37.4
4	2175	30.0	34	2148	46.1	64	2127	38.8
5	1730	31.9	35	2148	41.0	65	2119	81.9
6	1730	40.5	36	2148	48.7	66	2119	40.0
7	1730	40.1	37	2131	64.1	67	2119	37.1
8	1730	33.0	38	2131	49.1	68	2119	35.3
9	1764	37.3	39	2131	40.8	69	1708	74.8
10	1764	29.3	40	2131	51.9	70	1708	31.2
11	1764	36.4	41	1814	42.8	71	1708	46.2
12	1764	34.2	42	1814	27.7	72	1708	32.2
13	1778	32.9	43	1814	24.5	73	1727	35.5
14	1778	37.7	44	1814	25.0	74	1727	33.9
15	1778	51.0	45	2169	52.8	75	1727	26.8
16	1778	36.3	46	2169	37.2	76	1727	31.0
17	1739	31.6	47	2169	42.6	77	1743	37.4
18	1739	32.5	48	2169	47.2	78	1743	33.3
19	1739	32.4	49	2128	40.3	79	1743	34.7
20	1739	30.6	50	2128	42.6	80	1743	31.0
21	2159	39.8	51	2128	31.8	81	1734	41.3
22	2159	27.1	52	2128	39.6	82	1734	31.4
23	2159	40.0	53	2150	32.7	83	1734	30.7
24	2159	32.3	54	2150	61.4	84	1734	35.9
25	2136	44.2	55	2150	35.8	85	2160	81.1
26	2136	47.1	56	2150	37.2	86	2160	58.7
27	2136	38.4	57	2152	68.1	87	2160	46.5
28	2136	42.0	58	2152	40.4	88	2160	48.3
29	1737	33.5	59	2152	49.1	**Average**		**40.6**
30	1737	32.0	60	2152	37.5			

Table 8. Average deterioration of the oil for each of the twenty two buses for Company B

Bus #	% Deterioration	Bus #	% Deterioration
2175	32.55	1730	36.38
1764	34.30	1778	39.48
1739	31.78	2159	34.80
2136	42.93	1737	30.75
2148	55.00	2131	51.48
1814	30.00	2169	44.95
2128	38.58	2150	41.78
2152	48.78	2127	49.38
2119	48.58	1708	46.10
1727	31.80	1743	34.10
1734	34.83	2160	58.65

Table 9. R and MSE obtained for the neural models trained with dataset AB, using all the input variables and then just 12 selected input variables

# inputs	R (train set)	R (validation set)	R (test set)	R (all dataset)	MSE (all dataset)
21+2	0.97	0.84	0.86	0.93	0.27
10+2	0.96	0.86	0.84	0.94	0.15

Table 10. Confusion matrices of the errors of the model trained with data from AB

# inputs	Predicted	Companies A&B			Company A			Company B		
	3	1	0	**52**	0	0	**25**	1	0	**27**
21+2	2	7	**34**	3	4	**19**	2	3	**15**	1
	1	**77**	1	1	**36**	1	1	**41**	0	0
	3	0	0	**47**	0	0	**24**	0	0	**23**
10+2	2	4	**34**	7	2	**19**	4	2	**15**	3
	1	**81**	1	2	**38**	1	0	43	0	2
	Actual	1	2	3	1	2	3	1	2	3

7. Discussion and comparison of the results obtained

Using ANN models it was possible to determine companies A and B follow different policies, with Company B being more defensive than Company A. PCA confirmed these results, showing that Company B replaces the oil in the interval from 30 to 59 % of deterioration, while Company A sometimes even passes the limit established by the manufacturer. Using ANN it was

Table 11. Classification of the different samples by the human expert, the Artificial Neural Network, and level of deterioration of the oil according to PCA analysis. Samples where the human expert and the ANN differ are marked in bold.

Human	ANN	Det. (%)	Human	ANN	Det. (%)	Human	ANN	Det. (%)	Human	ANN	Det. (%)
1	1	31.7	1	1	51.6	2	3	37.5	1	1	36.3
1	1	13.6	2	2	54.8	3	3	64.6	1	1	31.6
2	2	49.9	1	1	35.7	3	3	56.7	1	1	32.5
1	1	18.8	1	1	31.7	2	2	37.4	1	1	32.4
1	1	52.1	1	1	13.6	2	2	38.8	1	1	30.6
3	3	150.6	2	2	49.9	3	3	81.9	3	3	39.8
1	1	35.6	1	1	18.8	3	3	40	2	2	27.1
1	1	17.6	1	1	52.1	2	2	37.1	2	2	40
1	1	13.4	3	3	150.6	2	2	35.3	2	2	32.3
3	3	79.1	1	1	35.6	1	1	74.8	3	3	44.2
1	1	64	1	1	17.6	1	1	31.2	3	3	47.1
2	2	83.6	1	1	13.4	1	1	46.2	3	3	38.4
2	2	73.6	3	3	79.1	1	1	32.2	3	3	42
1	1	46.5	1	1	64	**1**	**2**	35.5	1	1	33.5
1	1	54.1	2	2	83.6	1	1	33.9	1	1	32
2	2	69.2	2	2	69.2	1	1	26.8	1	1	29.1
3	3	162.1	3	3	162.1	1	1	31	1	1	28.4
3	3	128	3	3	128	**1**	**2**	37.4	3	3	84.2
3	3	88.3	3	3	88.3	1	1	33.3	3	3	46.1
3	3	118.3	3	3	118.3	1	1	34.7	2	2	41
1	1	48.3	1	1	48.3	1	1	31	3	3	48.7
1	1	45.6	1	1	45.6	**3**	**2**	41.3	3	3	64.1
2	2	63.7	2	2	63.7	1	1	31.4	**3**	**2**	49.1
2	2	55.2	2	2	55.2	1	1	30.7	1	1	40.8
3	3	66.7	3	3	66.7	1	1	35.9	3	3	51.9
1	1	32.5	1	1	32.5	3	3	81.1	1	1	42.8
1	1	54.2	1	1	54.2	3	3	58.7	1	1	27.7
3	3	89.1	3	3	89.1	3	3	46.5	1	1	24.5
3	**2**	**54.6**	**3**	**2**	54.6	3	3	48.3	1	1	25
1	1	51.7	1	1	51.7	1	1	31.1	2	2	52.8
3	3	69	3	3	69	1	1	30.4	2	2	37.2
2	2	70.7	2	2	70.7	2	2	38.7	**1**	**3**	42.6
3	3	73.1	3	3	73.1	1	1	30	3	3	47.2
3	3	85.3	3	3	85.3	1	1	31.9	3	3	40.3
1	1	49.3	1	1	49.3	1	1	40.5	**3**	**2**	42.6
1	1	55.8	1	1	55.8	1	1	40.1	2	2	31.8
2	2	54.3	2	2	54.3	1	1	33	3	3	39.6
2	2	62.3	**1**	**2**	62.3	1	1	37.3	1	1	32.7
3	**2**	72.5	**1**	**2**	72.5	1	1	29.3	3	3	61.4
2	**3**	73.3	3	3	73.3	1	1	36.4	2	2	35.8
1	1	31	1	1	31	1	1	34.2	2	2	37.2
3	3	86.8	2	2	73.6	1	1	32.9	3	3	68.1
2	**1**	31.7	1	1	46.5	1	1	37.7	3	3	40.4
3	3	66.4	1	1	54.1	1	1	51	3	3	49.1

possible to create predictive models to classify the oils, with a very high degree of accuracy. The predictions of the models are sometimes different from the decisions of the companies, but some of the errors may be due to poor decisions of the companies. Using PCA it was onfirmed that in both companies it is possible to identify oils which were replaced very early and oils which were replaced at a much more advanced degree of deterioration.

PCA also showed that the variables measured during oil analysis have different importance to assess the quality of the oil and predict the company decision. ANN modeling confirmed this result, since it was possible to train a model with very high accuracy using just 10 of the 21 variables.

Table 11 compares the results obtained for dataset AB. It shows the classification of the oil by the human experts, the classification given by the artificial neural network trained with 12 variables and the percentage of oil degradation calculated using PCA. The samples where the classification of the neural model is different from the human expert are marked in bold. As the table shows, there is a large variability of results. But in general the ANN's classification and the PCA classification are coherent and arguably better than the human experts. As a reference, the average average deterioration for class 1 is 37.86 % for the human experts and 37.03 % for the ANN. For class 2 it was 52.07 % for humans and 52.44 % for the ANN. And for class 3, the average deterioration is 73.05 % for humans and 74.14 % for the ANN. The average deteriorations show that the ANN was successful clustering the most deteriorated oils in class 3 and the less deteriorated oils in class 1, more than the human experts.

Looking in more detail at the situations where the ANN and the human classification differ, it is possible to conclude that there is a high probability that some of the misclassifications are human rather than design or training limitations of the ANN. For example, in Table 11 the first error happens in a sample where PCA determines a level of oil degradation of 54.6. That sample was classified by the company as a 3 and by the neural model as a 2. In fact, the average oil degradation for oils classified as 2 is approximately 52. The third error happens for a sample with oil degradation 31.7 according to PCA, which is below the average degradation for class 1 (approximately 37). That sample was classified as 2 by the experts and 1 by the neural network. Many of the remainder errors are similar to the ones already described, which shows the neural model is very much according to the results obtained with PCA analysis.

8. Main contributions

The present paper proposes different novel contributions to the state of the art, which can be highlighted as follows.

- The results show that it is possible to create good artificial neural models to classify the oils. Moreover, the models can perform possibly with even less errors than human experts.
- Using PCA, the relevance of the variables monitored for oil analysis was determined, thus providing a better insight into the importance of each variable.
- The results also show that a good neural model does not need to use all the variables. In fact, a good model was created with just 12 input variables. This helps the process of determining the right time for oil change.

9. Conclusion

Condition monitoring of engines' oil is very important to prolong the engine life, avoid unnecessary pollution and also accidents due to engine overheating or other failures. The present paper describes experiments to create different artificial neural models that can help classify the state of deterioration of the oils with high accuracy. Because of the different policies followed by different companies, it may be difficult to create one single model that fits all policies. But it was possible to create models that showed good performance for two different companies. Those models may even generalize and learn a balance between the two policies. The results of the neural models were convergent with the results of PCA. PCA determines which companies follow the best policies for oil replacement and which variables are best predictors. The present analysis may be useful to help companies make the best decisions at the best time, or even decide which variables are more important to monitor. Future research includes fine tuning the models with more data and proposing a model to automate the process, as well as testing other classification or future extraction techniques

References

1. Capone S, Zuppa M, Presicce D S, Francioso L, Casino F, Siciliano P. Metal oxide gas sensor array for the detection of diesel fuel in engine oil. Sensors and Actuators B: Chemical 2008; 131(1): 125-133, https://doi.org/10.1016/j.snb.2007.12.029.
2. Cerny B A, Kaiser H F. A study of a measure of sampling adequacy for factor-analytic correlation matrices. Multivariate Behavioral Research 1977; 12(1): 43-47, https://doi.org/10.1207/s15327906mbr1201_3.
3. Du L, Zhe J. As high throughput inductive pulse sensor for online oil debris monitoring. Tribology International 2011; 44(2): 175-179, https://doi.org/10.1016/j.triboint.2010.10.022.
4. El-Hag A H, Saker Y A, Shurrab I Y. Online oil condition monitoring using a partial discharge signal. IEEE Transactions on Power Delivery 2010; 26(2): 1288-1289, https://doi.org/10.1109/TPWRD.2010.2073551.
5. Gajewski J, Valis D. The determination of combustion engine condition and reliability using oil analysis by MLP and RBF neural networks. Tribology International 2017; 115: 557- 572, https://doi.org/10.1016/j.triboint.2017.06.032.
6. Ghobadian B, Rahimi H, Nikbakht A, Najafi G, Yusaf T. Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network. Renewable energy 2009; 34(4): 976-982, https://doi.org/10.1016/j.renene.2008.08.008.
7. Hongxiang T, Yuntao L, Xiangjun W. Application of neural network to diesel engine SOA. Third International Conference on Measuring Technology and Mechatronics Automation 2011; IEEE, https://doi.org/10.1109/ICMTMA.2011.141.
8. Kumar S, Mukherjee P, Mishra N. Online condition monitoring of engine oil. Industrial lubrication and tribology 2005; 57(6): 260-267, https://doi.org/10.1108/00368790510622362.
9. Li X, Li J, He D, Qu Y. Gear pitting fault diagnosis using raw acoustic emission signal based on deep learning. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (3): 403-410, https://doi.org/10.17531/ein.2019.3.6.
10. Li Y, Wang K. Modified convolutional neural network with global average pooling for intelligent fault diagnosis of industrial gearbox. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (1): 63-72, https://doi.org/10.17531/ein.2020.1.8.

11. Niu X, Yang C, Wang H, Wang Y. Investigation of ann and svm based on limited samples for performance and emissions prediction of a crdi-assisted marine diesel engine. Applied Thermal Engineering 2017; 111: 1353-1364, https://doi.org/10.1016/j.applthermaleng.2016.10.042.
12. Parlak A, Islamoglu Y, Yasar H, Egrisogut A. Application of artificial neural network to predict specific fuel consumption and exhaust temperature for a diesel engine. Applied Thermal Engineering 2006; 26(8-9): 824-828, https://doi.org/10.1016/j.applthermaleng.2005.10.006.
13. Raposo H, Farinha J T, Fonseca I, Galar D. Predicting condition based on oil analysis - a case study. Tribology International 2019; 135: 65-74, https://doi.org/10.1016/j.triboint.2019.01.041.
14. Rodrigues J, Costa I, Farinha J T, Mendes M, Margalho L. Modelling diesel engine oil condition using artificial neural networks. eMaintenance 2019, Stockholm, Sweden.
15. Shaban K, El-Hag A, Matveev A. A cascade of artificial neural networks to predict transformers oil parameters. IEEE Transactions on Dielectrics and Electrical Insulation 2009; 16(2): 516-523, https://doi.org/10.1109/TDEI.2009.4815187.
16. Westerholm R, Li H. A multivariate statistical analysis of fuel-related polycyclic aromatic hydrocarbon emissions from heavy-duty diesel vehicles. Environmental Science & Technology 1994; 28(5): 965-972, https://doi.org/10.1021/es00054a032.
17. Yonghui Y, Weihua W, Xinpin Y, Hanliang X, Chengtao W. An integrated online oil analysis method for condition monitoring. Measurement Science and Technology 2003; 14(11): 1973-1977, https://doi.org/10.1088/0957-0233/14/11/015.
18. Zhu J, He D, Bechhoefer E. Survey of lubrication oil condition monitoring, diagnostics, and prognostics techniques and systems. Journal of Chemical Science and Technology 2013; 2(3): 100-115.
19. Zhu J, Yoon J M, He D, Bechhoefer E. Online particle-contaminated lubrication oil condition monitoring and remaining useful life prediction for wind turbines. Wind Energy 2015; 18(6): 1131-1149, https://doi.org/10.1002/we.1746.
20. Zhu X, Zhong C, Zhe J. A high sensitivity wear debris sensor using ferrite cores for online oil condition monitoring. Measurement Science and Technology 2017; 28(7), https://doi.org/10.1088/1361-6501/aa6adb.

João RODRIGUES
CISE, Univ. Beira Interior, Covilhã, 6201-001, Portugal
and
Industrial Eng. and Management, Univ. Lusófona, Campo Grande 376, 1749-024, Lisboa, Portugal

Inês COSTA
Polytechnic Institute of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal

J. Torres FARINHA
Polytechnic Institute of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal
and
CEMMPRE, Coimbra University, DEM, Polo 2, 3030-290 Coimbra, Portugal

Mateus MENDES
Polytechnic Institute of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal
and
ISR, Coimbra University, DEEC, Polo 2, 3030-290 Coimbra, Portugal

Luís MARGALHO
Polytechnic Institute of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal

E-mails: j.antunesr@hotmail.com, a21260426@isec.pt, tfarinha@isec.pt, mmendes@isr.uc.pt, lmelo@isec.pt

Article citation info:
WIĘCŁAWSKI K, MĄCZAK J, SZCZUROWSKI K. Electric current as a source of information about control parameters of indirect injection fuel injector. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 449–454, http://dx.doi.org/10.17531/ein.2020.3.7.

Krzysztof WIĘCŁAWSKI
Jędrzej MĄCZAK
Krzysztof SZCZUROWSKI

ELECTRIC CURRENT AS A SOURCE OF INFORMATION ABOUT CONTROL PARAMETERS OF INDIRECT INJECTION FUEL INJECTOR

PRZEBIEG PRĄDOWY JAKO ŹRÓDŁO INFORMACJI O PARAMETRACH STEROWANIA WTRYSKIWACZEM PALIWOWYM WTRYSKU POŚREDNIEGO*

The article discusses results of the laboratory experiments in which fuel injectors used in indirect injection internal combustion engines were tested. During the experiments, numerous dosing cycles of the injectors were performed while changing the control parameters, due to which, the dosing characteristics were developed and influence of applied parameters on the resultant fuel flow determined. Simultaneously, the voltage and electric current waveforms in the injector coil were recorded, due to which finding links between the electric current characteristics and the determinants of the injector work was possible. The investigation has shown that parameters of electric current constitute a precise criterion for assessing the operation of the solenoid valve, because fuel flow is created due to the work ofelectric current. Thus, by observing the changes in the current flowing through the valve coil,it is possible to monitor precisely the correctness of the process of opening the flow and the electric current intensity, at which the flow began and to determine the mechanical quantities such as fuel dose and pressure. As a result, a characteristic is developed, that provides the links between the fuel pressure and the electric current at the point of lifting the needle, which is quite a novel approach. Such a characteristic can be used in diagnostics and control of fuel injectors as well as all kinds of electromagnetic valves.

Keywords: *injector, injector diagnostics, indirect injection, current waveform.*

Artykuł przedstawia wyniki eksperymentów laboratoryjnych polegających na testowaniu wtryskiwaczy paliwowych stosowanych w silnikach spalinowych z wtryskiem pośrednim. Podczas eksperymentów wykonano wiele cykli dawkowania wtryskiwaczy zmieniając parametry sterowania, dzięki czemu opracowano charakterystyki dawkowania i określono wpływ stosowanych parametrów sterowania na wynikowy przepływ paliwa. Jednocześnie rejestrowano przebiegi napięcia i natężenia prądu elektrycznego w cewce wtryskiwacza, dzięki czemu możliwe było powiązanie charakterystyk prądowych z determinantami pracy wtryskiwacza. Wykazano, iż parametry prądowe są precyzyjnym kryterium oceny pracy zaworu elektromagnetycznego, ponieważ dzięki wykonanej przez prąd pracy powstaje przepływ paliwa. Zatem poprzez obserwację zmian prądu płynącego przez cewkę zaworu, można precyzyjnie monitorować prawidłowość procesu otwierania przepływu oraz natężenie prądu, przy którym przepływ się rozpoczął oraz określać wielkości mechaniczne jak dawka i ciśnienie paliwa. Wynikiem badań jest opracowanie charakterystyki wiążącej ciśnienie paliwa z natężeniem prądu w punkcie podnoszenia iglicy, co jest podejściem nowatorskim. Taka charakterystyka może być wykorzystana w diagnostyce i sterowaniu wtryskiwaczy paliwowych oraz wszelkiego rodzaju zaworów elektromagnetycznych.

Słowa kluczowe: *wtryskiwacz, diagnostyka wtryskiwacza, wtrysk pośredni, przebieg prądowy.*

1. Introduction

The paper discusses results of the laboratory experiments, the purpose of which was finding links between the electric current parameters, controlling the work of the electromagnetic injector, and its mechanical parameters. Using the relation of mechanical parameters and the electric current quantities allows for supervision of the injector work and precise evaluation of the quality of individual injection phases, i.e. to control the technical condition of the fuel system and the injector during operation, all on the basis of the easily-observed current parameters.The knowledge of the relationship between these parameters can be applied not only in injector diagnostics but also to control injectors. This results from the fact that thanks to the work of the electric current, the fuel flow supplied to the injector under sufficient pressureis induced. The electric current flowing through the valve coil, generates the magnetic flux, and subsequently, the magnetic force acting on the needle (the element cutting the fuel off). This force, after overcoming all resistant forces, preventing this action enables lifting the needle. Additionally, the electric connection of the injector coil provides for both, its control and power supply. To summarise, the work of the magnetic force resulting from the flow of the electric current is a factor allowing for connecting the current with the mechanical quantities. Due to the accurate defining of all the above relations, the real starting and finishing points of successive phases of the fuel injection process can be determined, which, in turn, can be used in managing the engine work. Such identification may allow for developing a different strategy relative to the control logic commonly applied so far.An engine controller can be developed using continuously updated information about the real beginning andduration of individual injection phases. Such a controller will adequately correct the control of injectors, in case there were variations detected in starting and finishing of the dosing process, contrary to the methods used nowadays, that are based solely onmeasurements of the exhaust gases quality and corrections of injection parameters on the basis of the previously pro-

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

grammed maps [1]. The size of a fuel dose, determined on the basis of the current waveform integration, can serve as an additional verification of the control process,within a given control method of the flow of electric current through the injector coil.The profound knowledge of momentary parameters of injection phases may be used for precise injector diagnostics, too, therefore contributing to the early detection of failures that could affect the degradation of the exhaust gases purification system and even the engine damage. Thus, the early detection of damages in the fuel system is particularly vital.

Operation of a fuel injector may be evaluated on the basis of various quantities. Verified can be the control signal or the control result, i.e. the quality of the generated stream [3]. Shape, angle, and degree of the mixture atomization proves the quality of the dose obtained. Lee et al. [5], conducted an analysis of the dynamic behaviour of the solenoid valve while examining the phenomenon of electromagnetic field and flow through the exhaust channel [10]. In her work, Harantova [2]discussed a projectand analysis of the control and supply systems for injectors.Voltage of the injector power supply and pressure values in the fuel system at a given width of the control pulse (PWM - Pulse-Width Modulation)were analysed [9].At the same time, attempts were made, to achieve the highest possible efficiency and stability of the injector work. In [12], the method of fuel injector verification based on electric current characteristics and the possibility of correcting the operation of the injector by using appropriate control algorithms was presented. Tan et al. [14] shows how to adjust the injector control strategy to take into account changes in the resistance and inductance of its coil, taking into account the effects of aging. Nikolićet al. [8] discussed the processes occurring in the fuel supply systems, namely fuel injection, generation of the fuel mixture, combustion, and exhaust emissions. K function, as a correlation between the speed of the fuel flow and fuel pressure depending on its type [4] [11] was determined. Fluctuations in the fuel pressure have a major impact on the atomization, combustion and size of the fuel stream, as well as on delays in the injector operation [6] [7]. Stępień [13] described the process of sludge formation in the injector fuel channels, the impact of the resulting sludge on changes in diagnostic parameters indicating the degree of degradation of the injector. In [4], a novel method was discussed, using Coriolis flowmeters (CFM), and a new, patented technique of signal processing for measuring the fuel flow. Measuring the flow speed of individual injectors of the engine in real time was proved possible, which enables the accurate assessment of the injection process. Merola et al. [7] presented a method of verifying fuel injection and the combustion process using optical diagnostics, using an endoscopic system coupled with a CCD camera mounted in the intake manifold.

The abovementioned works are examples of current publications regarding methods of evaluation of the fuel injector operation. It is a final element of the fuel system and undisturbed engine work depends highly on its flawless functioning. Defects of the injector may lead to degradation of the exhaust purification systems (catalytic converter). This is why the appropriate diagnostics and control of the fuel injector is a particularly important operation-related issue. In the following sections, an innovative diagnostic method forfuel injectors and all kinds of electromagnetic valvesis discussed. This method is based on identification of values of the intensity of electric current controlling the injectorat the point of lifting the needle. The value of the current at this point is not a constant. It is dependent upon several factors resulting fromthe injector's properties and control parameters. Accuracy of determination of the values of the electric current and the detailed analysis of the determinants for injector control allows forusing the correlation between the value of the electric current at the characteristic points of the current-related waveform of the dosing injector and its mechanical parameters. On the basis of observation of the characteristic current-related points, practically all electric and mechanical failures of the injector can be detected, with no need of OBD-based diagnostics or other methods, including those based on removing the injector from the vehicle engine. The method of diagnostics presented in this article may be applied during operation without removing the injector from the engine, and after appropriate implementation in the engine controller, an automatic tool is obtained, that will ensure an early detection of damage in both, the injector and the fuel system. In the subject literature regarding the diagnostics of fuel injectors, there are no methods mentioned, that are based on the correlation described in Section 3 of this article.

2. Dose and fuel flux

The electric current parameters precisely describe the phenomena they affect. This is illustrated in Fig. 1, where the increasing surface area under the current waveform (continuous red line) reflects the ever-increasing flow of the specific mediumin response to the prolonged voltage impulse from 2 ms to 15 ms. In each of the waveforms in fig. 1, at 0.6 V there is a dotted dark blue line, denoting the photodetector reading, transforming the laser light into the voltage. In the tests performed at the test bench, the stream of fuel disturbed the laser light running under the injector nozzle (decreased voltage output by approx. 0.15 V), illustrating the real fuelflow.

The injector work is based on an undisturbed cooperation of the electric and hydraulic systems, controlling of which is supervised by the electronic system. The current flow during the preset injection time is defined by the continuity equation:

$$\nabla\ j + \frac{\partial \rho}{\partial t} = 0 \tag{1}$$

where: j – electric current density,

$\nabla j = \frac{\partial j}{\partial x} + \frac{\partial j}{\partial y} + \frac{\partial j}{\partial z}$ – scalar multiplication with the nabla vector operator,

ρ – density of electric charge,

t – time [s].

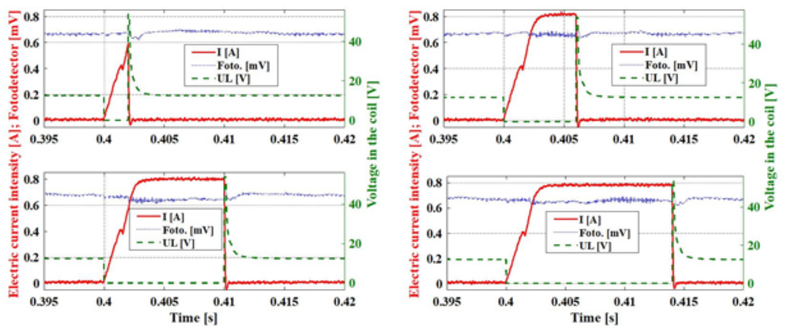


Fig. 1 Waveforms of fluctuations of electric current, voltage, and intensity of laser light corresponding to the increased injection times, from 2 ms to 15 msregistered at the test bench

The electric current density$j$is a result of differentiation of the electric current relative to the surface where it flows:

$$j = \frac{dI}{dA} \tag{2}$$

where: I – electric current [A], A – cross sectional area [m^2].

The motion of the injector needle results from the change in the magnetic flux φ, which in turn is a result of the electric current I flowing through the coil. This is described as a flow of the electric charge$Q$:

$$Q = \int_{t_1}^{t_2} i(t)\,dt = \frac{1}{RQ}\int_0^{\varphi} \frac{d\varphi}{dt}\,dt = \frac{\varphi}{R} \tag{3}$$

where: R isresistance [Ω]and φ -magnetic flux [Wb].

Due to the flow of electric charges, the fuel flow is obtained, described by analogical equation. The stream continuity equation:

$$\rho * \nabla v + \frac{\partial \rho}{\partial t} = 0 \tag{4}$$

In this formula, v denotes volume $\left[m^3\right]$.

Changes in flowing fuel flux may be correlated with the electric charge flowing during injector's dosing action. Using the current-related waveforms allowed for generating the characteristics of the dose and of the fuel flux, mapping the electric current waveforms. For instance, at the injection time of $10\,ms$ and the injection pressure $p = 0.3\,\text{MPa}$, the obtained stream flux amounted to: $0.0061 \pm 0.000124\frac{\text{mg}}{\text{ms}}$, which is related to the flow of the electric charge $Q = 0.00668\,\text{C}$. The electric charge Q was computed by integrating the current-based waveform (3). Subsequently, the theoretical volume of the fuel stream was calculated in accordance with the equation (5):

$$\dot{m} = f_{s-c} * A * \rho * \sqrt{\frac{2*(p_1 - p_2)}{\rho_i}} \tag{5}$$

where: f_{s-c} – flow coefficient,

ρ_i – density of fuel during flow through the injector nozzle,

p_1 – pressure of fuel before the injector,

p_2 – pressure of fuel after the injector.

In the example shown above $(10\,\text{ms} \text{and} 0.3\,\text{MPa})$, the obtained mass flow amounted to: $0.00556\frac{\text{mg}}{\text{ms}}$. The result is burdened with a measurement error resulting from the uncertainty of parameter measurements from equation (5) such as: fuel pressure or fuel density. The result of the calculation differs from the actual by $0.000416\frac{\text{mg}}{\text{ms}}$. Because of the uncertainty of the measurement of the quantity from equation (5), it is a satisfactory result.

Fig. 2 showscurrent waveforms for different voltage pulse widths, from 1.6 ms to 10 ms, superimposed on each other.

An electric charge $q_{1,6ms} = 0.000418\,\text{C}$ has been assigned to the 1.6 ms range (first range from the left in Fig. 2, indicted by a golden continuous line), while range of 10 ms (the whole of the current waveform – marked with the red continuous line) has an electric charge of the value $q_{10ms} = 0.00668\,\text{C}$. The difference between $q_{1,6ms}$ and q_{10ms} is 16–fold. The areas of successive ranges as well as electric charges are proportional to their corresponding values.

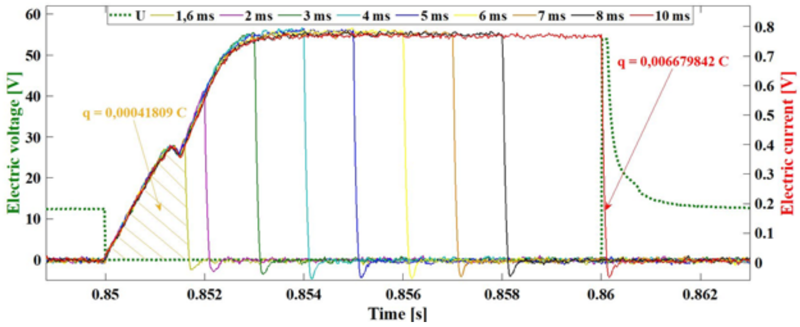


Fig. 2. Current waveforms for different voltage pulse widths, from 1.6 ms to 10 ms, superimposed on each other, with the assignment of the area under the waveform

In the case of the current ranges for various fuel pressure values (Fig. 3), the differences between electric charges representing successive current waveforms are too small to assess (on the abovementioned basis) at which value of the fuel pressure the dose was generated and what its size was. These values lie within therange of measurement uncertainty. The size of the dose and stream can be related to the value of the fuel pressure, defined on the basis of electric current value at the point of the needle lifting, which will be the main focus of the second part of this article.

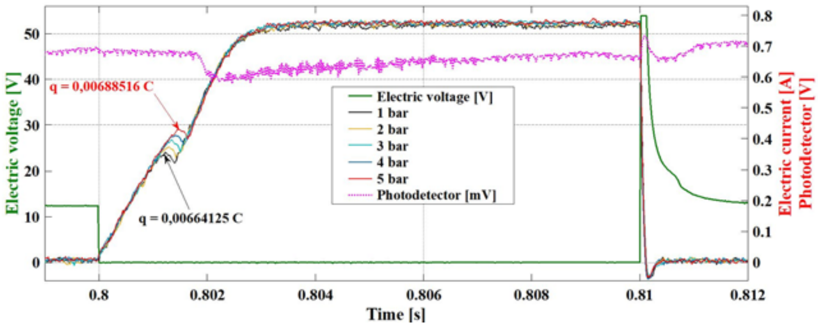


Fig. 3. Current waveforms for increasing values of fuel pressure, from 0.1 to 0.5 MPa

3. Fuel pressure

The current waveform recorded during generation of a single fuel dose, with a preset eight-millisecond length of a control pulse is shown in Fig. 4. The electric circuit of the injector coil consists of the source of the electromotive force (ε_0), resistance (R), and inductance (L) (circuit RL). Increase in the current in the RL circuit is described by the Kirchhoff's equation (6), defining the shape and values of the electric current varying over time. In the case of the injector's current waveform, as a result of the work by the magnetic force, i.e. lifting of the needle, equations describing the electric current must take into consideration the resistance overcome by this force, which is discussed in this section of the article.

$$I_{op} = \frac{\varepsilon_0}{R} * \left(1 - e^{\frac{R}{L}}\right) \tag{6}$$

The shape of the waveforms observed during the work of the electromagnetic valve depends not only on the electric current parameters (current intensity, voltage) and geometry of the valve core. Location of the characteristic points (both value and the time lapse) depends on the specific density of the flowing medium and value of the pressure that reaches the valve.In this part of the article reference is made to bench tests in which the pressure downstream of the injector was constant. Therefore, the results refer only to the pressure before the injector. The purpose of the analysis is to show the trend of changes in electric current at the needle lifting point depending on the pressure (in fact, the pressure difference before and after the injector). During the engine operation of the injector, the pressure in the intake manifold changes, which change must be taken into account in the analyzes performed.

The fuel flow into the injector inlet channel totally changes the time-related voltage-current waveform. As a result of the fuel density and pressure action (it is density, in fact, since pressure causes its increase), a greater magnetic force needs to be generated in order to lift the needle. At the specific value of current at the point where the needle is lifted (I_{op}), a corresponding pressure value of the fuel before it reaches the valve needle, can be precisely matched $\left(p_{inj}\right)$. Fig. 4 shows two injector current-related waveforms, for two different fuel pressure values before the injector. Change in the injection pressure from 0.2 MPa to 0.8 MPa results in a change in the current at the point where the needle is lifted, from 0.379 A to 0.495 A.

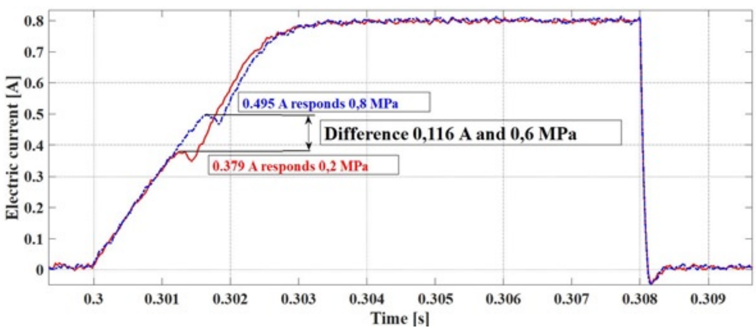


Fig. 4. Graphic representation of the relationship I_{op} (p_{inj})

On the basis of the laboratory research, for each injector type, the function of relationship of the current intensity (at the point where the needle is lifted) and the injection pressure can be determined:

$$I_{op} = f\left(p_{inj}\right) \tag{7}$$

The fuel flow can be obtained after the magnetic force F_m has overcome all the forces counteracting the needle lifting:

F_p – force resulting from the fuel pressure,

F_i – inertial force,

F_f – frictional force,

F_s – force of the spring.

$$F_m > F_p + F_i + F_f + F_s \tag{8}$$

Magnetic force F_m is a derivative of the energy originating in the coil as a result of the current. The magnetic force sufficient to lift the needle is defined by means of the current intensity, measured at this point of the current-related waveform of the injector I_{op}. The most significant resistant force is force F_p, the remaining elements are constant or vary based on this force's change. The consequence of it being that the magnetic force F_m is a function of the current intensity I_{op}. And the current intensity I_{op} is a function of the force resulting from the fuel pressure p_{inj}:

$$F_m = f\left(I_{op}\right) \tag{9}$$

where: F_m – magnetic force resulting from the magnetic flux [N],
I_{op} – current intensity at the point of the needle lifting [A],
p_{inj} – fuel injection pressure [MPa].

For the increasing fuel pressure, the force necessary to lift the needle grows, thus the current required to generate the magnetic flux of the adequate value, increases, too:

$$I_{op1} = f\left(p_1\right) < I_{op9} = f\left(p_9\right)\left\{p_1 < p_9\right\} \tag{10}$$

Relationships (10) are graphically represented in Fig.5. It shows an image of the current waveform of injector with the increase in current intensity at the point of the needle lifting marked, depending on the injection pressure.

Current intensity at the point of needle lifting, (Fig.5) corresponds closely to the injection pressure. Moreover, it is unrelated to the injection time, it depends on the value of generated magnetic force. The greater the difference between the pressure before the injector and the pressure in the intake manifold (after the injector), the higher the current I_{op}. The mapping of such a characteristic in the relation (6) required using the coefficient f_{press}. After inserting the coefficient into the differential equation, in the component $\frac{\varepsilon_0}{R}$ (6), the expected value of the current I_{op} will be obtained (11):

$$I_{op} = \left(f_{press} * \frac{\varepsilon_0}{R}\right) * \left(1 - e^{\frac{R}{L}}\right) \tag{11}$$

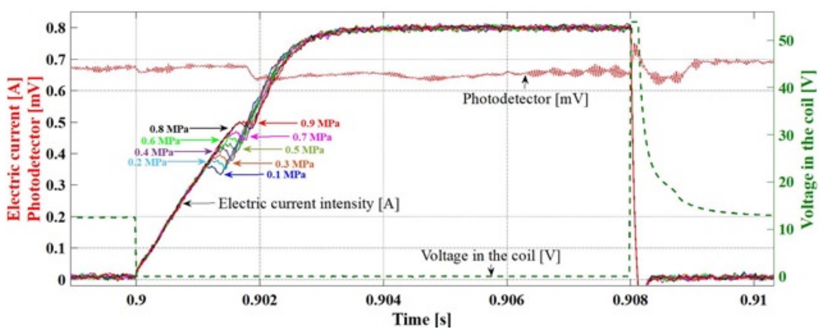


Fig. 5. Changes in electric current intensity at the point of the injector needle lifting at different injection pressure values

The first component determines the maximal value that will be reached I_{op}, whereas the second component affects the exponential tending to this value (Kirchhoff's equation). Due to the expression (11), the resistant forces in the current-defining equation in the circuit of the working injector (circuit RL) are taken into consideration, which can be observed in the equations modelling the current ranges in the previous section. Characteristic I_{op} (p_{inj}), necessary in the detailed modelling of the injector current waveform can be utilised to evaluate the given injector's technical state, in comparison to different types of injectors or in the early diagnosing of its faults. All kinds of defects of the electromagnetic valve will result in change in the current intensity at the point of the needle lifting. The injector can be monitored during its work, controlling this parameter and the value of injection pressure from the pressure sensor. Characteristic $I_{op}\ (\ p_{inj})$ is a relationship of the linear character. As the pressure keeps growing, a little increase in the proportion of the current intensity I_{op} to p_{inj} can be observed. Characteristics $I_{op}\left(p_{inj}\right)$ have been developed for several different types of injectors with different degree of wear. For a given type of injector, characteristics have similar character but they are not identical, which results from the differences in their operation. The characteristic of the injectors "a" and "b" (Fig.6) changes throughout the whole range of the injection pressure (from 0.1 to 0.8 MPa).

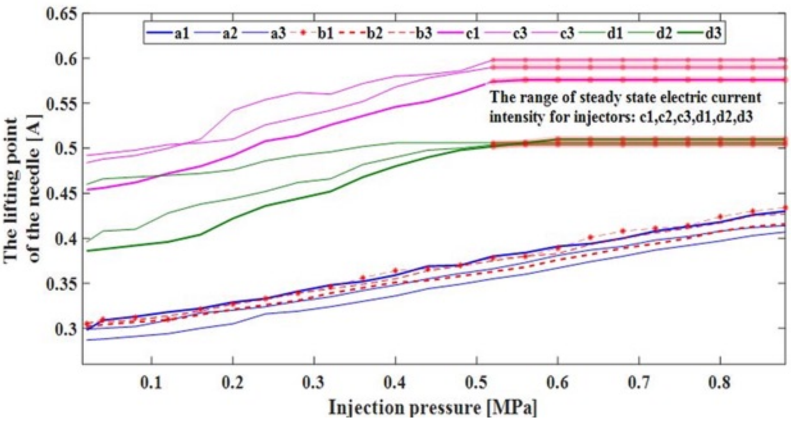


Fig. 6. Characteristics $\boldsymbol{I_{op} = f\left(p_{inj}\right)}$ *for injectors: "a", "b", "c" and "d"*

The magnetic flux resulting from the flowing current, generated around the coils of these injectors is greater than the generated by the coils of the injectors "c" and "d" (Fig.6). The current intensity for injectors "c" and "d" from pressure of 0.5 MPa transitions into the steady state (maximal – horizontal line). Above this pressure value, they do not increase their doses and the current intensity at the point of the needle lifting is constant (such as the current intensity of the steady state in the circuit). Ability to overcome the forces counteracting the needle lifting at the lower current I_{op} indicates a given injector's greater efficiency.

The characteristic of the relationship between the electric current at the point of the needle lifting and the mechanical resistances of the lifting action results from the electrical and geometric propertiesof the injector, fuel system, and preset control parameters. A precise determination of the corresponding parameters allows for verification of the technical state of the fuel system, also in the real time. This innovative approach may be applied both in diagnostics and to introduce corrections in injector control. The presented analysis shows that on the basis of the selected points of the time-related current-voltage waveform, observed during injector dosing, such parameters of injector control as the injection pressure can be determined, and combining this information with the surface area below the current intensity waveform, allows for characterising the mass flow.

4. Summary

The article discusses results of the laboratory experiments testing fuel injectors used in internal combustion engines with the indirect injectionwithout turbo boost. In the course of experiments, waveforms of voltage and current in the injector coil were recorded. Also, the dosage characteristics were developed; the influence of applied control parameterson the resultant fuel flow was determined. On the basis of the above, it was indicated that the current-related parameters constitute a precise criterion for the injector work evaluation. Correlation of the current waveforms with dosage parameters of the fuel injectors enabled the conclusion that fluctuations within the waveforms result not only from the current-related properties. They also depend on mechanical properties affecting the flow (4), such as density, injection pressure, or resistance of the needle motion. The conclusion can be made that the shape of the waveform of the electric current (6)in the circuit of the injector coilcontains information about mechanical parameters of injection (11), as well as electric and mechanical properties of the injector. The range of the current waveform or the flowing electric charge (3) can be correlated with the fuel flux(5). In order to map the fuel dose precisely, it is necessary to determine at what pressure value occurred the fuel flow. The characteristic point of the current waveform, used in the discussed analysis. The value of the current at this point relates directly to the forces counteracting the lifting action (8). The largest of these is the force resulting from the fuel pressure, which is one of the variables. The second variable is the pressure downstream of the injector, i.e. the pressure in the intake manifold. The reference to the current at the needle lifting point makes sense under the same comparison conditions, i.e. with current pressure control downstream of the injector.

A conclusion can be made that a radical change in the electric current at this point takes place following the change in the fuel pressure (7), which has been proved by the performed laboratory experiments.

Observation of the electric current waveforms and analysis thereof, in combination with the previously generated characteristic of correlated parameters, allows for verification of the correctness of the fuel dosing process, and for rendition of the assessment of technical state of the fuel system and the injector. The presented current waveforms may be monitored by means of the controller in the real time, during operation, due to which such verification can support the OBD and the control systems for the exhaust emissions, allowing for a faster failure detection in the fuel system. The relation between the value of the electric current at the point of needle lifting and the fuel pressure may be used not only in fuel injector diagnostics but also to verify different kinds of electromagnetic valves.

The results of laboratory experiments presented in this work reflect a new viewpoint regarding the fuel injector diagnostics. An additional diagnostics discussed here may be used in the real time during operation of the injector or electromagnetic valve. The „on-line" diagnostics can contribute to an earlier detection of damages, thus protecting the exhaust purification system from degradation and even from engine damage. The function of the fuel pressure can be also used to control the injectors. Determination of the real injection phase may help managing the engine work. This means the change of strategy relative to the one used so far. The presented information could be applied in the module controlling engine work, using in the control algorithm the information on the real moment of starting and finishing the process of dosing, and not be functioning only on the basis of the previously developed algorithms, corrected through the adaptation of control and the lambda probe.

Acknowledgement

Project financed by the National Centre for Research and Development under the DOB-BIO9-04-02-2018 programme.

References

1. Ashok B, Ashok D, Ramesh S, Kumar C. A review on control system architecture of a SI engine management system. Elsevier Science. Annual Reviews in Control 2016; 41: 94–118, https://doi.org/10.1016/j.arcontrol.2016.04.005.
2. Harantová V, Otáhalová Z, Kasanický M. Estimation of fuel consumption based on data from opening fuel injector valve. Transportation Research Procedia 2019; 40: 236–243, https://doi.org//10.1016/j.trpro.2019.07.036
3. Kay P J, Bowen P, Bowen P, Gold M R. Sapsford S M. Studies of gasoline direct-injection sprays at elevated ambient gas temperatures and pressures. Atomization and Sprays 2012; 4 (22): 305–331, https://doi.org/10.1615 / AtomizSpr.2012005411.
4. Leach F, Davy M H, Manus P H, Tombs M, Zhou F. A new method for measuring fuel flow in an individual injection in real time. SAE International Journal of Engines 2018; 11 (6): 687–696, https://doi.org//10.2307/26649124.
5. Lee G S, Sung H J, Kim H C. Multiphysics analysis of a linear control solenoid valve. Journal of Fluids Engineering 2013; 135(1): 011104–0111041, https://doi.org//10.1115/1.4023079.
6. Li S, Nehl T, Gopalakrishnan S, Omekanda A, Namuduri C, Prasad R. A Dynamic solenoid model for fuel injectors. IEEE Energy Conversion Congress and Exposition. Portland, USA 2018; 3206–3213, https://doi.org//10.1109/ECCE.2018.8558394.
7. Merola S S, Tornatore C, Sementa P. Optical investigation of the fuel injector influence in a PFI spark ignition engine for two-wheel vehicles. Journal of Mechanical Science and Technology 2012; 26(1): 223–233, https://doi.org//10.1007/s12206-011-0801-5.
8. Nikolić B, Jovanović M, Milošević M, Milanović S. Function K as a link between fuel flow velocity and fuel pressure, depending on the type of fuel. Facta Universitatis, Series: Mechanical Engineering 2019; 15(1): 119–132, https://doi.org//10.22190/FUME160628003
9. Peng M, Deng X, Liang X. Control strategy calibration of fuel injection impulse width on EFI motorcycle engine. 2011 International Conference on Electric Information and Control Engineering. Wuhan, China 2011; 2047–2053, https://doi.org//10.1109/ICEICE.2011.5776821.
10. Pielecha I, Skowron M, Mazanek A. Evaluation of the injectors operational wear process based on optical fuel spray analysis. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2018; 20(1): 83–89, https://doi.org//10.17531/ein.2018.1.11.
11. Qiu T, Songa X, Lei Y, Dai H, Cao C, Xu H, Feng X . Effect of back pressure on nozzle inner flow in fuel injector. Fuel 2016; 173: 79–89, https://doi.org//10.1016/j.fuel.2016.01.044.
12. Sebok M, Jurcik J, Gutten M, Kornciak D, Roj J, Zukowski P. Diagnostics and measurement of the gasoline engines injection system. Przeglad Elektrotechniczny 2015; 1(8): 77–80, https://doi.org//10.15199/48.2015.08.20.
13. Stępień Z. A study of factors influencing the formation of harmful deposits in the diesel engine injectors. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 9(3): 331–337, https://doi.org//10.17531/ein.2017.3.3.
14. Tan T, Park J S, Leteinturier P. (2016). Enhanced Injector Dead Time Compensation by Current Feedback. SAE 2016 World Congress and Exhibition. Detroit, United States 2016; 2016-01–0088: p.8, https://doi.org//10.4271/2016-01-0088.

Krzysztof WIĘCŁAWSKI
Jędrzej MĄCZAK
Krzysztof SZCZUROWSKI

Faculty of Automotive and Construction Machinery Engineering
Warsaw University of Technology
ul. Narbutta 84, 02-524Warszawa, Poland

E-mails: krzysztof.wieclawski@pw.edu.pl, jedrzej.maczak@pw.edu.pl, krzysztof.szczurowski@pw.edu.pl

Article citation info:
YAN S, MA B, WANG X, CHEN J, ZHENG C. Maintenance policy for oil-lubricated systems with oil analysis data. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 455–464, http://dx.doi.org/10.17531/ein.2020.3.8.

Shufa YAN
Biao MA
Xu WANG
Jianhua CHEN
Changsong ZHENG

MAINTENANCE POLICY FOR OIL-LUBRICATED SYSTEMS WITH OIL ANALYSIS DATA

POLITYKA UTRZYMANIA RUCHU UKŁADÓW SMAROWANYCH OLEJEM W OPARCIU O DANE Z ANALIZY OLEJU

Maintenance of oil-lubricated systems plays a significant role in reducing unexpected system failures and improving machine availability. This paper deals with the oil-lubricated systems subject to gradual degradation that revealed by metal wear debris monitored using oil analysis. Oil-lubricated systems usually undertake several preventive maintenances during operation, after each maintenance, the system typical restores to an intermediate state between good-as-new state and bad-as-old state due to system aging such as cumulative wear. Furthermore, oil-lubricated systems often operate continuously in mission execution with availability constraints. However, existing literature still lacks a method to integrate the availability constraints with the system aging into the cause of optimizing the maintenance policy. To fill this gap, this paper develops a maintenance policy optimization method to determine the optimal maintenance threshold joint considering the availability constraints and the system aging. A case study of the power-shift steering transmission systems modelled by a wiener process is presented to illustrate the proposed method in practical application.

Keywords: *oil-lubricated system; maintenance policy; oil analysis; availability constraint; system aging.*

Konserwacja układów smarowanych olejem odgrywa istotną rolę w eksploatacji maszyn, umożliwiając zmniejszenie liczby nieoczekiwanych uszkodzeń i poprawiając dostępność maszyn. Niniejszy artykuł dotyczy układów smarowanych olejem ulegających stopniowej degradacji, którą można mierzyć za pomocą analizy oleju, monitorując zawartość drobin metalu powstających na skutek zużycia. Podczas swojej pracy, układy smarowane olejem zwykle poddawane są kilkakrotnie przeglądom zapobiegawczym. Po każdej konserwacji, układ wraca do stanu pośredniego między stanem fabrycznej nowości (as good as new) a stanem "jak przed konserwacją" (as bad as old), co wynika ze starzenia się systemu, m.in. skumulowanego zużycia. Co więcej, systemy smarowane olejem często działają w sposób ciągły, wykonując misje z ograniczeniami dostępności. Jednak w istniejącej literaturze wciąż brakuje metody, która pozwalałaby na zintegrowane ujęcie ograniczeń dostępności i starzenia w celu optymalizacji polityki utrzymania ruchu. Aby wypełnić tę lukę, w niniejszym artykule opracowano metodę optymalizacji polityki utrzymania ruchu, dzięki której można określić optymalny próg konserwacji z uwzględnieniem zarówno ograniczeń dostępności jak i starzenia się systemu. Możliwość praktycznego zastosowania zaproponowanej metody zilustrowano na podstawie studium przypadku układów przekładni kierowniczych zamodelowanych za pomocą procesu Wienera.

Słowa kluczowe: *układ smarowany olejem; polityka utrzymania ruchu; analiza oleju; ograniczenie dostępności; starzenie się systemu.*

Acronyms

CBM: condition-based maintenance
LCM: lubricant condition monitoring
MLE: maximum likelihood estimation
PCA: principal component analysis
PHM: prognostics and health management
PM: preventive maintenance
PSST: power-shift steering transmission

1. Introduction

Lubricant condition monitoring (LCM) has been widely used in oil-lubricated machinery benefit from the advance in analytical instruments and monitoring techniques, see, e.g., [2, 8, 14, 16, 22, 24] and references therein. The analysis results of oil samples not only give an indicator of the suitability of the lubricating oil for continued use but also provide information about the wear condition of oil-lubricated machines [20, 31]. This advantage enables the oil analysis technique to become one of the most popular data acquisition techniques used in condition monitoring, prognostics, and maintenance. For a comprehensive review of the use of oil analysis data for maintenance decision support, readers are referred to [28] and the references therein. In these existing studies, many efforts have been made in establishing a reasonable degradation model or estimating an accurate residual life. However, the existing literature still lacks an effective method to utilize oil analysis data to implement optimal maintenance to ensure the operation safety, availability and economic gains of the oil-lubricated systems. Therefore, this paper aims to fill this literature gap by addressing the condition-based maintenance (CBM) problem for the oil-lubricated systems with selected oil analysis data.

Spectral oil analysis is one of the most widely used oil analysis techniques to observe the increasing trends of wear debris concentrations in the oil-lubricated systems [9]. It can be used to identify the severity of wear conditions and even the impending system failure without dismantling the machine [24,30]. Specifically, if the oil spectral data reaches a predetermined failure threshold, the oil-lubricated system is considered as failed that should be maintained or repaired [19,31]. As shown in Fig. 1, for example (Liu et al., [15]), the oil spectral data, Cu, is used to monitoring the degradation of a Power-Shift Steering Transmission (PSST) system; the PSST system experiences degradation during operation and eventually fails (unable to operation satisfactorily) and needs to be replaced when the spectral oil data, $X(t)$, reaches the threshold, D_F , at time t_F . If the degradation profile of the oil-lubricated system can be modelled and evaluated, the preventive maintenance (PM) action might be implemented based on the collected oil spectral data before the system failed.

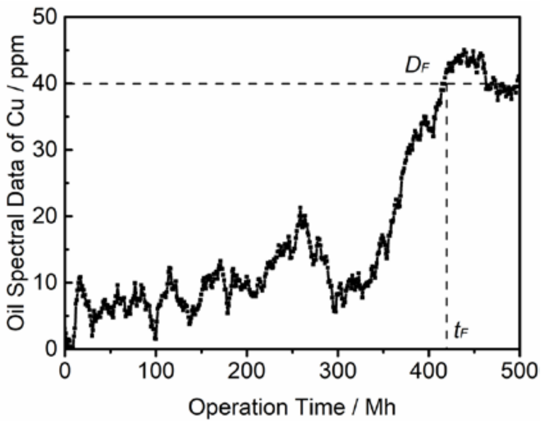


Fig. 1. Condition monitoring using spectral oil data

Accordingly, to achieve an optimal implementation of the PM actions on oil-lubricated systems, there are two essential requirements must be addressed: (1) a reliable degradation model that can characterize the degradation profile of oil-lubricated systems, such that the degradation condition can be accurately estimated and provided as an input for PM; and (2) a reasonable maintenance objective function that can optimize the PM actions of oil-lubricated systems, such that the operational requirements of the maintained system can be better satisfied.

Over the past few years, many research efforts have been made on the requirement (1). Many types of degradation models, such as proportional hazard models [6,36], hidden Markov models [5,29], and Wiener process models [26,31], have been developed in modelling the degradation evolution of the oil-lubricated systems. Among these models, the Wiener process model can provide a good description of the system degradation process by describing the degradation as a non-monotonic process. Vališ et al. [27] considered using the wiener process with the selected spectral oil data (i.e., Fe, and Pb) to predict the expected moment when a soft system failure occurred and proved that the spectral oil data could provide a useful reference for CBM. In this paper, it is assumed that the degradation process follows a normal distribution by regarding the degradation as an increment of a large amount of small wear. Yan et al. [31] proposed a Wiener process-based degradation model for assessing the material wear of the friction pairs in an oil-lubricated system, and in [34] further investigated the influences of the spectral measurement uncertainty in modelling the system degradation. It is known that the collected oil spectral data usually shows a non-monotonic trend because of measurement uncertainty. Therefore, in this paper, the degradation profile of oil-lubricated systems is described using a Wiener process.

With the degradation profile predicted, PM (e.g., dismantling inspection, lubricating oil replacement) will be carried out. While extensive work has been done in developing these degradation models, the current literature still falls short in addressing the requirement (2), i.e., how to construct a reasonable maintenance objective function for the oil-lubricated systems. Most of the existing studies [32, 33, 35] simply use a cost minimization function as the objective, and the maintenance costs (e.g., oil analysis cost, oil replacement cost, dismantling inspection cost and system failure cost) are minimized in a full life cycle. However, for some critical systems in military equipment, mining machinery, and power industry, whose failure is critically hazardous and often leads to catastrophic consequences, and therefore, using a cost minimization objective in these cases may be problematic [1, 12]. In these critical systems, the ratio of time on the operation (uptime) compared with the time on the maintenance (downtime) is more important than the cost [3, 7]. Thus, in this paper, a more practical objective function, the maximum availability, is utilized to measure maintenance effectiveness.

Since the degradation of the oil-lubricated system is described by a wiener-process-based model, and the maintenance objective is constructed as an availability function, the PM threshold is set as the decision variable forming the optimal maintenance policy. In order to optimize the policy, the effect of each PM action should also be addressed. In the literature of CBM, the effect is usually divided into perfect PM and imperfect PM. However, the existing research concerning the oil-lubricated systems [32, 33, 35] simply assumes that the PM action is perfect without considering the system aging, and the maintained system will be recovered to a good-as-new state. Actually, using a perfect PM assumption for the oil-lubricated systems is problematic, as discussed in many related studies [4,10]. For instance, as discussed in Yan et al. [31], "because of the system residual damage such as cumulative wear, the degradation will start at some non-zero value after each PM and randomly increasing with the order of PM cycle." This phenomenon is called the system aging property that will shorten the time interval to the next PM. To address this issue, in this paper, the PM action is assumed imperfect and can only partially restore the system, i.e., recover the system to an intermediate state between bad-as-old state and good-as-new state.

Motivated by the above observations, in this paper, an optimal PM problem is considered in terms of the availability requirement considering the system aging property. The main objective of this paper is to develop a maintenance decision method for oil-lubricated systems based on oil analysis data in order that the PM threshold can be achieved with maximum availability. Compared with the existing studies in Liu et al. [35] and Yan et al. [32], the proposed method has the following innovations: (1) considering the system availability requirement as the objective to PM policy optimization; and (2) involving the system aging property in the optimization model. To be specific, the oil-lubricated systems that monitored under oil spectral analysis, with an availability requirement subject to periodic unperfect PM (e.g., lubricating oil replacement), are considered. For modelling the PM process, a wiener process is utilized for the degradation model, and a geometric process is used for the system residual damage after imperfect PM actions. In addition, the availability requirement under the short-run availability constraint is adopted in the optimal PM policy model. Finally, a case study for several PSST systems in military vehicles is provided to illustrate the proposed method. The results show the practicality, effectiveness and robustness of the proposed method.

The framework of the proposed maintenance policy optimization method is shown in Fig. 2. The remainder of the paper is structured as follows. Section 2 gives the motivation of the concerned optimal PM problem for oil-lubricated systems and its properties under multiple imperfect PM action. The system aging property description, the operation time model, and the maintenance duration model are introduced in Section 3. Section 4 gives the formulation of the optimization problem for the oil-lubricated system and the corresponding calculation procedures for achieving the optimal PM policy. In Section 5, a case study for several PSST systems is provided for illustration of the proposed method. Finally, conclusions and discussions are drawn in Section 6.

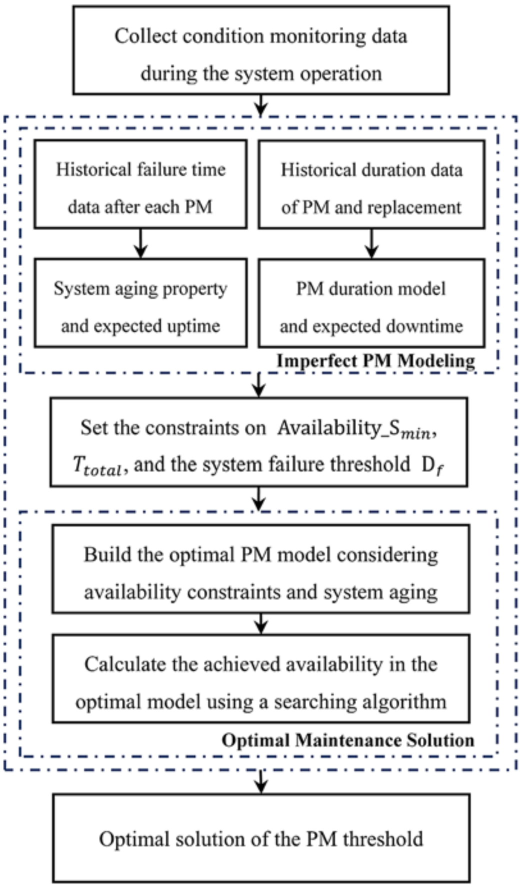


Fig. 2. Framework for the maintenance policy optimization

2. PM policy for oil-lubricated systems

This section provides a general description of the concerned optimal PM problem for oil-lubricated systems. This paper deals with the oil-lubricated systems that degrade over time in mission execution conditions, and oil spectral analysis is periodically conducted to monitor the degradation severity of the system during the whole lifecycle (from an initial to failure). As shown in Fig. 3, the degradation process $\{X(t),\ t \ge 0\}$ that measured using collected spectral oil data is used to predict the degradation profile of the oil-lubricated system. When the degradation profile of the oil-lubricated system reaches a predetermined PM threshold, D_{PM}, at the end of a mission, PM action will be implemented immediately before the next mission (e.g., at time $R_{i=1,2,3}$ in Fig. 3), and the time for the PM actions is M_i. Alternatively, the oil-lubricated system will continue to operate with no PM action until the next oil inspection. Please note that the PM actions mainly include the lubricating oil replacement, the possible dismantling inspection, and the replacement of the potential components, usually are imperfect, i.e., the PM actions cannot recover the oil-lubricated system to the good-as-new state. When a new one replaces the used oil-lubricated system (e.g., at time $t_0^{'}$ in Fig. 3), the system will be fully restored to a good-as-new state and a time ζ æ is incurred.

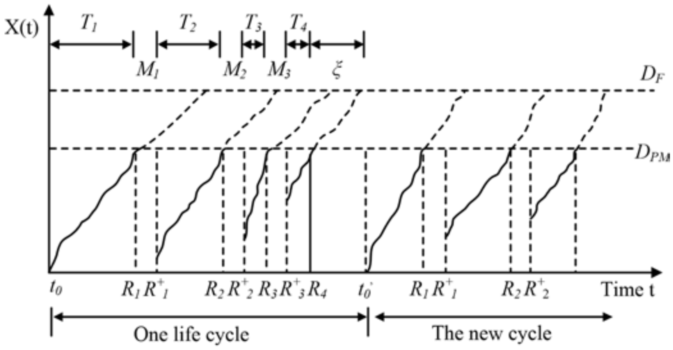


Fig. 3. State transition of an oil-lubricated system with PM and system replacement

Let T_i ($i = 1, 2, \ldots$) represents the operating time of the oil-lubricated system after a PM action. In reality, the expected uptime period, $E[T_i]$, for the maintained oil-lubricated system to reach the PM threshold in a cycle show a decreasing trend because of the imperfect nature of PM and the system aging such as cumulative wear. On the other hand, the expected downtime period, $E[M_i]$, for the maintained oil-lubricated system to perform each PM action in a cycle usually show an increasing trend, since a more extended period is required to maintain a severely degraded system than a healthier one. However, for the oil-lubricated systems used in military equipment, mining machines, and power industries, the mission interruption time is strictly limited. As a result, the purpose of this research is to find the optimal PM policy to maximize the mission availability of an oil-lubricated system, that is, to maximize the expected uptime compared with the expected downtime.

Based on the above description of the concerned optimal PM problem, the mission availability of the oil-lubricated system shows a decreasing trend in a replacement cycle, since the expected downtime increases whereas the expected uptime decreases. Therefore, the mission availability metric in Eq. (1) is adopted as the optimization objective to assess the effectiveness of the maintenance policy. The mission availability for oil-lubricated systems in a life cycle is defined as:

$$Availavility\_\mathrm{L} = \frac{(\text{expected uptime})/\text{ cycle}}{(\text{expected uptime} + \text{expected downtime})/\text{ cycle}} \quad (1)$$

where the "cycle" is defined as the period between two consecutive replacements of the oil-lubricated systems, the "downtime" includes the time for all PM actions and system replacement. Besides, for oil-lubricated systems, frequent PM actions are not allowed in engineering practice. Thus, the oil-lubricated systems must satisfy a minimum short-run availability. Specifically, the average short-run availability metric defined in (2) is adopted to represent the mission constraint of an oil-lubricated system after the i th PM action, within a replacement life cycle, as described above.

$$Availavility\_\mathrm{S}(i) = \frac{\textit{expected uptime after ith PM}}{(\textit{expected uptime} + \textit{expected downtime})\textit{after ith PM}} \quad (2)$$

As many researchers have investigated, a system will no longer be suitable for operation when the defined average short-run availability decreases to a certain level [4,10,31]. Thus, whenever the short-run availability, $Availavility\_\mathrm{S}_{min}$, is lower than a threshold after the N th PM action, a system replacement will be conducted (e.g., at time R_4 in Fig. 3) and, the oil-lubricated system would be restored to a good-as-new state. Specifically, the replacement policy for the oil-lubricated system is summarized as:

$$\begin{cases} Availavility\_S(i) > Availavility\_S_{min} & i \in [1, N-1] \rightarrow \text{PM actions} \\ Availavility\_S(N) \leq Availavility\_S_{min} & \rightarrow \text{System replacement} \end{cases} \quad (3)$$

Remark 1: The objective of this paper is to determine the optimal value of the PM threshold, D_{PM}, to maximize the achieved availability defined in Eq. (1) for a required service time under the constraint of the average short-run availability defined in Eq. (2). We refer to this newly proposed policy as *availability limit policy*. Unlike the existing *cost limit policy* [25-27], where the maintenance cost is assumed as the optimization objective, the proposed policy considering the mission availability to initiate the PM actions and system replacement. In this way, the achieved availability of the maintained oil-lubricated system in the whole mission execution can be maximized, which is a practical advantage for the maintenance of oil-lubricated systems used in critical systems, such as the PSST system in armored tank vehicles [31], the wind turbine gearbox portion of wind power systems [23], and the diesel engine in marine freighters [11].

Equation (3) can be numerically calculated using a searching algorithm. From the above description, it is conculcated that the optimal maintenance policy is heavily dependent on the system aging property, and the expected uptime and downtime, which can be fitted by the historical time data of system operation. Thus, in the following, the residual damage, and operation time model and maintenance time model are investigated. The aim is to accurately initiate the PM policy optimization model and obtain the optimal PM threshold.

3. Imperfect PM model considering system aging

In order to establish the considered maintenance optimization model, it is necessary to investigate the influence of the aging property of oil-lubricated systems. According to the system aging characteristics described in Section 2, the system aging property, the required operation time model and PM duration model involving multiple PM actions are given as follows.

3.1. The System Aging Property

Recall that the PM actions would not restore the maintained oil-lubricated system to a good-as-new state but an intermediate state due to the system aging, i.e., cumulative wear of friction pairs. Thus, the characteristics of the aging propriety after a single PM action and the relationship with multiple PM actions are investigated for the maintained oil-lubricated system.

In general, the PM actions will at least restore the maintained system to a healthier condition than the state before the PM. Fig. 4 provides a schematic diagram of a viable imperfect PM model to describe the PM effect considering the system aging. That is, the residual damage, $X\left(R_i^+\right)$, of the maintained system after each PM action in a replacement cycle falls randomly in the interval [0, D_{PM}].

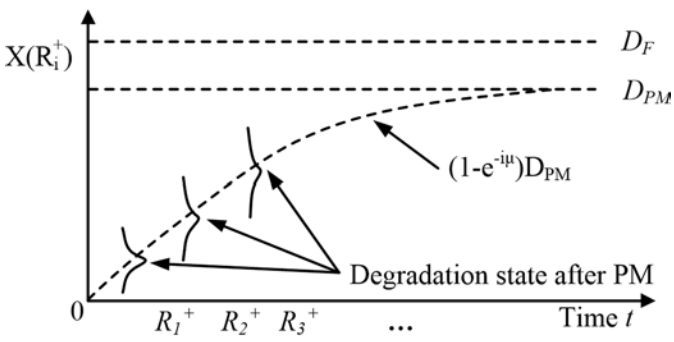


Fig. 4. System aging property with imperfect PM

It is commonly assumed that the residual damage in a replacement cycle follows a stochastic process with an exponentially increasing mean and a constant variance. Actually, such an assumption is widely used in modeling the aging property of the degrading systems, as shown in many applications [3, 4, 10]. To be specific, the mean and variance are given as follows:

$$\begin{cases} E\left[\dfrac{X\left(R_i^+\right)}{D_{PM}}\right] = 1 - \exp(-iu) & (4) \\ Var\left[\dfrac{X\left(R_i^+\right)}{D_{PM}}\right] = \sigma^2 & (5) \end{cases}$$

where $u \geq 0$ is a constant value describing the effort of the PM actions, $\sigma^2 \geq 0$ is a constant value describing the variance of the PM actions.

Remark 2: The system aging property is represented by the proposed residual damage model, which can be understood in the following way. If [$u > 0,\ \sigma^2 > 0$] is assumed, the residual damage, $X\left(R_i^+\right)$, becomes a random variable constrained in a finite interval [0, D_{PM}]. In this case, the probabilistic models with compact support including the *Beta* distribution are commonly used to describe the system aging property [13]. When [$u > 0,\ \sigma^2 > 0$] is satisfied, the maintenance effect results in restoring the system health condition depending on the selection of D_{PM}. On the other hand, [$u = 0,\ \sigma^2 = 0$] refers to the perfect maintenance that recovers the maintained system back to the good-as-new state. Obviously, such a perfect PM assumption is a special case of the proposed system residual damage model, as have used in the existing literature [33, 32].

3.2. The operation time model

Since the system aging property of the maintained oil-lubricated system is determined by using the residual damage model, the expectation of the operating time after the $(i+1)$th PM action, T_{i+1}, can be calculated with a given D_{PM}, namely:

$$E\left[T_{i+1}\right] = E\left[E\left[T_{i+1} \mid X\left(R_i^+\right)\right]\right] = \int_0^{D_{PM}} E\left[T_{i+1} \mid X\left(R_i^+\right)\right] f_{X\left(R_i^+\right)}(X)\,\mathrm{d}x \quad (6)$$

where $f_{X\left(R_i^+\right)}$ is the probability density function (PDF) of the residual damage, $X\left(R_i^+\right)$. Recall that probabilistic models with compact support are commonly used to describe the residual damage. Thus, in this paper, the *Beta* distribution is employed to depict the residual damage, $X\left(R_i^+\right)$, just like many researchers have been done [3, 4, 7, 10, 13]. Such that the PDF series, $f_{X\left(R_i^+\right)}(X)$, is then defined as:

$$f_{X\left(R_i^+\right)}(X) = \frac{1}{D_{PM}} \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} \left(\frac{x}{D_{PM}}\right)^{\alpha_i - 1} \left(1 - \frac{x}{D_{PM}}\right)^{\beta_i - 1} I_{\{0 \leq x \leq D_{PM}\}} \quad (7)$$

where $i \in [1, N]$ represents the number of PM actions; the relationship between model parameters $[\alpha_i > 0, \beta_i > 0]$ and [u, σ^2] is as follows:

$$\begin{cases} E\left[\dfrac{X\left(R_i^+\right)}{D_{PM}}\right] = \dfrac{\alpha_i}{\alpha_i+\beta_i} = 1-\exp\left(-iu\right) & (8) \\ Var\left[\dfrac{X\left(R_i^+\right)}{D_{PM}}\right] = \dfrac{\alpha_i\beta_i}{\left(\alpha_i+\beta_i\right)^2\left(\alpha_i+\beta_i+1\right)} = \sigma^2 & (9) \end{cases}$$

The parameters, α_i and β_i, can be approximately obtained based on the statistical analysis with the historical failure time data after each PM action and, the parameters, u and σ^2, in the residual damage model can then be easily estimated by using the Maximum Likelihood Estimation (MLE) method. For detailed procedures of the parameter estimation, readers are referred to the literature [17] and the references therein.

Remark 3: It can be easily proved that the mean of the system residual damage in a replacement cycle is showing an increasing trend with the sequence of PM action, as shown in Eq. (4) and (8). That is, the ability of the maintained oil-lubricated system to operate in a healthy condition is weakening, which is in line with reality.

3.3. The maintenance time model

Recall that a more extend PM duration may be required with the system operation dues to the possible dismantling inspection and the replacement of the potential components [27, 31] since a long time is required to maintain a severely aged system than a slighted aged one. Thus, in this paper, it is assumed that the PM durations shows an increasing trend in a replacement cycle.

To be specific, let us denote the duration needed to implement the i th PM action as M_i, and is independent with M_j for any $i \neq j$. Similar to the choice in existing literature [10, 13], it is assumed that the PM durations are exponentially distributed, described as:

$$E\left[M_i\right] = \gamma_0 D_{PM}\exp\left(i\gamma_1 D_{PM}\right) \quad (10)$$

where $\gamma_0 > 0$ and $\gamma_1 \geq 0$ are constants. In reality, γ_0 and γ_1 can be easily estimated by using the real field data of the PM durations to fit the exponential distribution in Eq. (10).

Remark 4: The commonly used PM duration model in existing literature [4,10] is included in the proposed model as its special case when $\gamma_1 = 0$ is assumed. It can be clearly concluded from Eq. (10) that $E\left[M_i\right] \geq E\left[M_j\right]$ for any $i > j \geq 1$, which is in line with reality.

Remark 5: The rationality of such modeling assumptions described above is verified in the case study. Please note that such assumptions may not be suitable for all applications, and some other distributions assumption like Weibull distribution may be used in other applications.

4. PM policy formulation and optimization

The oil-lubricated system used in critical systems is usually constrained with mission availability. As mentioned earlier, the PM actions would recover the maintained oil-lubricated system to an intermediate state due to the system aging, i.e., cumulative wear of friction pairs, which results in a more extend PM duration and a shorter expectation of the operating time. Thus, with the residual damage model developed in Section 3, the goal of this section is to develop an availability maximum PM policy for the oil-lubricated system with the average short-run availability constraint.

4.1. The optimization problem formulation

For the oil-lubricated systems monitored using regular oil analysis, the achieved availability can be maximized by controlling the PM threshold. After each system replacement, the oil-lubricated system will be recovered to a good-as-new state. Therefore, the considered optimal PM problem is calculated by the following programming formulation:

$$\text{MAX Availability\_L}\left(D_{PM}\right) = \frac{\sum_{i=1}^{N+1} T_i}{\sum_{i=1}^{N+1} T_i + \sum_{i=1}^{N} M_i + \zeta} \quad (11)$$

$$\text{Subject to} \begin{cases} 0 < D_{PM} \leq D_F \\ Availability\_S\left(i\right) \geq Availability\_S_{min};\ i \in \left[1, N\right] \\ E\left[\sum_{i=1}^{N+1} T_i\right] \geq T_{total} \end{cases}$$

The second constraint ensures that the average short-run availability is not lower than a threshold. If this constraint elapsed, the oil-lubricated system will be replaced by a new one. The concerned average short-run availability is expressed as $\text{Availability\_S}(\text{i}) = T_{i+1} / \left(T_{i+1} + M_i\right)$. The third constraint indicates that the total running time of the oil-lubricated system should not be less than a predetermined period of time, T_{total}, as the obtained optimal PM policy is meaningless when the total running time is shorter than a threshold.

4.2. The optimization algorithm description

An optimization algorithm is designed to solve the programming formulation to obtain the optimum PM threshold. Specifically, the optimal PM threshold is calculated by using a searching algorithm to search over the range (0, D_F] that maximizes the achieved Availability_L. The optimization algorithm is as follows.

Step 1: Initialize the PM threshold D_{PM} within the range (0, D_F] with a small value;

Step 2: Calculate the expected operating time, $E\left[T_{i+1}\right]$, using Eu. (10), and the expected PM duration, $E\left[M_i\right]$, using Eq. (6);

Step 3: Calculate the average short-run availability, $\text{Availability\_S}(\text{i}) = T_{i+1} / \left(T_{i+1} + M_i\right)$, for every PM action until the constraint, $\text{Availability\_S}_{min}$, is firstly violated, and set the obtained N as the number of PM actions;

Step 4: Calculated the total running time, $\sum_{i=1}^{N+1} T_i$; if $\sum_{i=1}^{N+1} T_i \geq T_{total}$ is satisfied, calculated the achieved availability Availability_L for the current PM threshold D_{PM}, otherwise, go to next;

Step 5: Adjust the PM threshold D_{PM} by a small increment unless the PM threshold $D_{PM} > D_F$ is satisfied and repeat steps 2-4.

Step 6: Choose the optimal D_{PM} with the maximization of the achieved availability Availability_L.

5. Case study

In this section, a practical case study for PSST systems used in large engineering machinery is provided to illustrate the detailed application procedures of the proposed PM policy optimization method and to investigate the effectiveness of the proposed method. The PSST system usually works in severe work conditions with constraints on mission availability, especially for military usage like armored tank vehicles. Recent research shows that about 50% of in-service failures of the PSST systems result from metal wear debris [34,35]. As such, the concentrations of metal wear debris are used to reveal the degradation of the PSST systems, and the oil spectral analysis is used to reveal the metal wear debris in the lubricating oil [14]. For more than 10 years, we have collected the oil field data of the PSST system for more than one thousand samples, of which each dataset contains the time data concerning the system operating, the PM and system replacement, as well as the spectral oil data that represents the degradation evolution of the PSST system. A detailed description of the sampling and analysis processes can be found in [31]. In this way, the optimal PM threshold with required availability constraints can be obtained by using the proposed maintenance policy optimization method.

5.1. Development of the degradation model

Using the above mentioned spectral oil data, the degradation profile of the PSST system can be established and, the expectation of the operating time can then be predicted with the PM threshold using Eq. (6). Since the collected spectral oil data shows increasing but not necessarily a monotonic trend; thus, in this paper, a Wiener process-based degradation model is adopted. To be specific, let $X(t)$ represents the system degradation condition at time t. The degradation is expressed by:

$$X(t)=\theta t+\tau B(t), \quad \theta \geq 0 \tag{12}$$

where $B(t)$ is a standard Brownian movement, $\tau B(t) \sim N\left(0,\tau^2 t\right)$, $\tau > 0$ represents the diffusion coefficient; $\theta \geq 0$ is the drift coefficient that characterizes the degradation rate of the monitored oil-lubricated system.

According to the CBM theory, the system residual life is defined based on the first hitting time (FHT) of the Wiener process-based degradation model $\{X(t), t \geq 0\}$. Given the system failure threshold D_f, such FHT can be defined by $T_f = \inf\left\{t \mid X(t) \geq D_f\right\}$ [15]. T_f has an inverse Gauss distribution and the corresponding PDF and cumulative distribution function (CDF) is:

$$\begin{cases} f_{T_f}(t)=\dfrac{D_f}{\sqrt{2\pi t^3 \tau^2}} exp\left(-\dfrac{\left(D_f-\theta t\right)^2}{2\tau^2 t}\right) & (13) \\ F_{T_f}(t)=\Phi\left(\dfrac{-D_f+\theta t}{\tau\sqrt{t}}\right)+exp\left(\dfrac{2\theta D_f}{\tau^2}\right)\Phi\left(\dfrac{-D_f+\theta t}{\tau\sqrt{t}}\right) & (14) \end{cases}$$

Based on the homogeneous Markov property and the independent increment property of the Wiener process, the residual life T_i after the i th PM with residual damage $X\left(R_i^+\right) < D_f$, can be formulated as:

$$T_i = \inf\left\{t \mid X(t)+X\left(R_i^+\right) \geq D_f\right\} \text{ if } X\left(R_i^+\right) < D_f\text{; otherwise } T_{X\left(R_i^+\right)} = 0$$

Similar to the FHT distribution in Eq. (13) and (14), the residual life T_i knowing $X\left(R_i^+\right)$ also conforms to an inverse Gauss distribution. Thus, with the residual damage of the system estimated, the condition PDF and CDF of the residual life, T_i, after the i th PM actions can be easily obtained by replacing D_f by $D_f - X\left(R_i^+\right)$ in Eq. (13) and (14) as:

$$\begin{cases} f_{T_i}\left(t|X\left(R_i^+\right)\right)=\dfrac{D_{PM}-X\left(R_i^+\right)}{\sqrt{2\pi t^3\tau^2}}\exp\left(-\dfrac{\left(D_{PM}-X\left(R_i^+\right)-\theta t\right)^2}{2\tau^2 t}\right) \\ F_{T_i}\left(t \mid X\left(R_i^+\right)\right)=\Phi\left(\dfrac{-\left(D_f-X\left(R_i^+\right)\right)+\theta t}{\tau\sqrt{t}}\right)+exp\left(\dfrac{2\theta\left(D_f-X\left(R_i^+\right)\right)}{\tau^2}\right)\Phi\left(\dfrac{-\left(D_f-X\left(R_i^+\right)\right)+\theta t}{\tau\sqrt{t}}\right) \end{cases}$$

(15) (16)

For the PSST system we used here, the wear debris of Cu shows the most contribution to the system failure, according to our previous research in [14]. Therefore, in this paper, we focus on wear debris of Cu (as shown in Fig. 1) to illustrate the proposed method. When the spectral oil data of Cu reaches the failure threshold, $D_F = 0.04‰$, the system is considered failed. Using the spectral oil data with the above-mentioned degradation modeling method, the degradation profile of the monitored PSST system can be established, and the system residual life can then be estimated. Specifically, the parameters $\left[\theta,\tau^2\right]$ of the degradation model were estimated using the MLE method. A detailed description of the MLE method can be found in [17]. As such, the degradation model parameters are $\theta = 3.15\times10^{-5}$, $\tau^2 = 9.45\times10^{-6}$, which will be used to initialize the optimization problem.

Remark 6: Please note that not all the mental wear debris shows the same degradation pattern. Some other types of wear debris may be used in other machines, such as the wear debris of Fe for marine diesel engines [30]. The selection choice of wear debris can be made based on correlation analysis, principal component analysis (PCA), and clustering analysis, as mentioned in [27].

5.2. Estimation of the imperfect PM model

Recall that the PM actions (e.g., dismantling inspection, lubricating oil replacement) do not restore the oil-lubricated system to a good-as-new state but an intermediate state due to the system aging property such as cumulative wear. As a result, the ability of the PM actions to keep the system operates in a healthy condition is weakening. In other words, the system aging property will shorten the expected operating time and prolong the PM duration. Thus, in order to establish a realistic optimization model for the considered PSST system, the expected operating time and the PM durations of the considered PSST system are estimated according to the description in Section 4.

5.2.1. The operating time model

The operating time model is estimated using the operating time data collected during the mission. A preliminary analysis shows that the time period in service between the PM actions shows a decreasing trend, i.e., $E\left[T_i\right] > E\left[T_j\right]$ for $i < j$. Such orders following the sequences of PM actions describe the relative degree of the system aging by the current PM action to the previous ones, and lead to the decisions on PM implementation and, eventually, the system replacement. Thus, the historical failure time data following the sequences of PM actions are used to estimate the parameters [u, σ^2] in the system residual damage model according to the proposed method in Section 3.1. As such, the model parameters are $u = 0.672$, $\sigma^2 = 0.008$, which represents the effort and the variance of the PM actions.

Remark 7: Please note that the determination method for the system residual damage is an open issue in the existing literature, especially for the complex systems with multiple components, which is not the research focus of this paper. Of course, other distribution such as Weibull distribution can also be used according to the system aging property of the monitored system. For detailed methods of system residual damage estimation, readers are referred to the literature [18] and the references therein.

5.2.2. The maintenance time model

The PM duration model is estimated using the maintenance duration data collected during the mission. A preliminary analysis including Anderson-Darling test [21] shows that the PM durations in a replacement cycle shows an exponentially increasing trend, i.e., $E\left[M_i\right] \geq E\left[M_j\right]$ for $i > j$, which is in line with our proposed maintenance time model. Therefore, we estimated the parameters using the collected field data of the PM durations to fit the exponential distribution in Eq. (9). As such, the model parameters are $\gamma_0 = 1.54$, $\gamma_1 = 0.357$.

Remark 8: It is noted that a more extend PM duration is required with the system operating. That because a longer time is required to maintain an aged system than a slighted aged one due to the possible dismantling inspection and potential components replacement. Of course, other models, such as a linear decreasing model, can also be used according to the PM durations of the monitored system.

5.3. Solution of the optimal PM policy

With the estimated expectations of the operation time and PM durations, the established PM policy optimization model can be initialized, and the optimal PM threshold can be eventually obtained. For the sake of illustration, the parameters of the optimization model are shown in Tab. 1.

Table 1. Parameters of the optimization model

Parameters	Values	Parameters	Values
D_f	0.04‰	u	0.672
$Availability\_S_{min}$	0.6	σ^2	0.08
T_{total}	60(Day)	ζ	3.5(Day)
θ	3.15×10^{-5}	γ_0	1.54
τ^2	9.45×10^{-6}	γ_1	0.357

To be specific, the proposed PM optimization problem can be understood in the following way: The degradation of the monitored PSST system is monitored using regular oil spectral analysis and modeled using a Wiener process with the parameters $\theta = 3.15\times10^{-5}$, $\tau^2 = 9.45\times10^{-6}$. When the collected spectral oil data reaches the threshold $D_F = 0.04‰$, the system is defined as failed. The system is replaced based on system failure or the short-run availability $Availavility_{S_{min}} = 0.6$ cannot be sustained, and a time $\zeta = 3.5$ (Day) is required. On the other hand, the system is preventively maintained upon a threshold D_{PM} with lubricating oil replacement and possible dismantling inspection and components replacement [35]. The effects of the PM actions are characterized by $u = 0.672$, $\sigma^2 = 0.008$, (see Eq. (4) and (5)). And the PM durations are characterized by $\gamma_0 = 1.54$ and $\gamma_1 = 0.357$, according to Eq. (10).

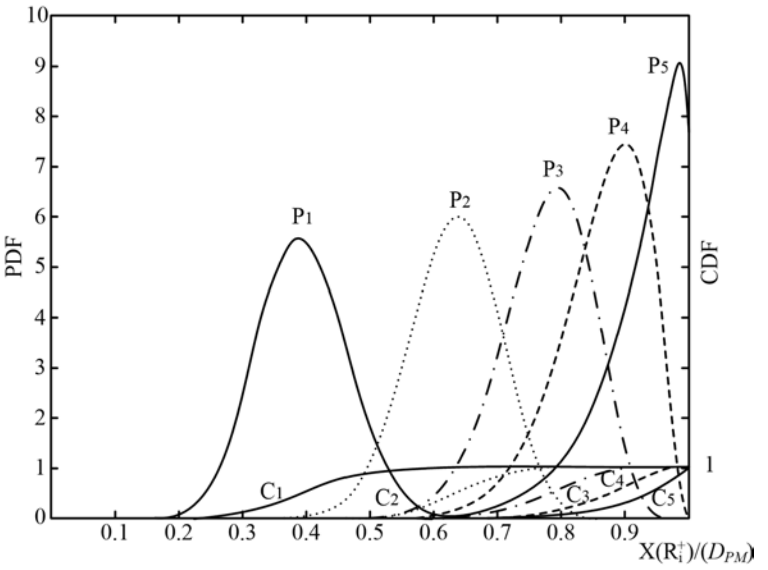


Fig. 5. System residual damage after PM actions

According to Eq. (4) and (5), the system residual damage, $X\left(R_i^+\right), i = 1,2,\ldots$, is shown in Fig. 5 with the number of PM actions increase. It can be seen that the estimated system residual damage shows an increase with the rise of the number of PM actions. When the number of the PM actions reaches 5, the system residual damage will rise to the PM threshold, D_{PM}. It is noted that frequent PM actions will be performed if the system residual damage approaches the PM threshold, D_{PM}, too much. This phenomenon will lead to a violation of the short-run availability constraint. So, the PM threshold should be optimized to satisfy the mission requirement.

Since the threshold, D_{PM}, is a decision variable in engineering practice as well as our proposed method; the objective is to determine the optimal PM threshold. Above all, the problem is formulated as:

MAX Availability_L(D_PM)

$$\text{Subject to} \begin{cases} 0 < D_{PM} \leq 0.04‰ \\ Availability\_S(i) \geq 0.6; \quad i \in [1, N] \\ E\left[\sum_{i=1}^{N+1} T_i\right] \geq 60 \end{cases}$$

After calculation, the average short-run availability after multiple PM actions with various D_{PM} as well as the number of PM actions is

Table 2. Short-run availability $Availability\_S$ with different PM threshold for the PSST system

Number of PM actions	D_{PM} (‰)									
	0.004	0.008	0.012	0.016	0.02	0.024	0.028	0.032	.0.036	0.04
0	Total service time < T_{total}				1	1	1	1	1	1
1					0.973	0.973	0.973	0.973	0.973	0.973
2					0.947	0.943	0.938	0.932	0.925	0.918
3					0.875	0.868	0.856	0.842	0.829	**0.817**
4					0.766	0.750	**0.722**	**0.687**	**0.646**	0.589
5					**0.658**	**0.632**	0.576	0.525	0.448	0.397
Total service time (Day)					65.7	70.2	74.7	79.4	85.5	92.4

presented in Tab. 2. In addition, the total service time is presented for different PM threshold, D_{PM} , with an increment of $0.004‰$. The cases when the PM threshold, D_{PM} , less than $0.02‰$ are excluded for that the corresponding total service time is violated with the required time constraint, $T_{total} \geq 60$ (Day).

In Tab. 2, the bold font indicates the smallest number of PM actions, N , when the short-run availability constraint is violated. The evolution of the achieved availability, Availability_L , with the increase of PM threshold, D_{PM} , is shown in Fig. 6. The optimal maintenance policy with the maximized long-run availability is finally obtained as:

$D^*_{PM} = 0.024‰$, $N^* = 5$, and the associated long-run availability,

$Availability\_L^* = 0.876$.

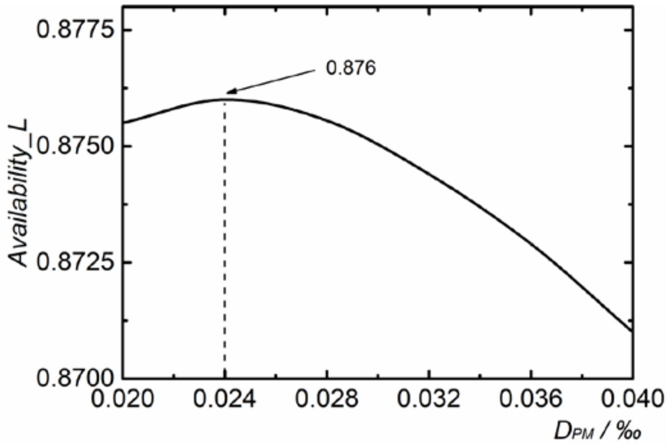


Fig. 6. The achieved availability with PM threshold

Table 3. Sensitive analysis of the model parameters on the optimal threshold

Variation	ζ		γ_0		γ_1		u		σ^2	
	-10%	10%	-10%	10%	-10%	10%	-10%	10%	-10%	10%
Optimal threshold D^*_{PM} (‰)	0.024	0.024	**0.028**	0.024	**0.028**	0.024	0.024	**0.028**	0.024	0.024

Furthermore, it can be seen from Fig. 6 that, with the PM threshold, D_{PM} , increasing, the achieved availability may increase even though the short-run availability is decreasing (e.g., at the PM threshold $D_{PM} = 0.02‰$ and $0.024‰$ in Fig. 6). This phenomenon may result from the balance of the system uptime expectation and downtime expectation. Compare with the PM threshold, $D^{*'}_{PM} = 0.02‰$ that determined empirically [35], the optimal PM threshold in our paper can obtain a higher achieved availability and a longer service time.

Since the parameters, u , σ^2, ζ, γ_0 and γ_1, describing the optimization model are estimated based on the population-wide characteristics from historical operation data, the variations of such parameters for a particular system may affect the optimization results. Thus, in order to analyze the influence of the parameters to the optimal result, a sensitivity analysis of the model parameters of the PM durations and replacement time to the PM policy is investigated, as shown in Tab. 3. It can be seen that the optimal PM threshold is slightly sensitive to the parameter variations of the PM duration model (e.g., γ_0 and γ_1), and the operating time model (e.g., u), and is insensitive to the system replacement time (e.g., ζ). As such, it can be concluded that the system aging property and the PM durations description have important influences on the planning of the maintenance policy. Therefore, we need to consider the accurate record of the information concerning system operation and maintenance.

6. Conclusions and discussions

In this paper, a maintenance policy optimization method is proposed for oil-lubricated systems based on the selected oil analysis data. Most of the existing maintenance policies focused on minimizing the maintenance cost without considering the mission constraints and the system aging property. Compared with these existing methods, the proposed method considers the system aging property, as well as the mission constraints such as the short-run availability, such that the achieved system long-run availability of the maintained system can be maximized and therefore, constitute the novelty of this paper. The proposed method was verified for several PSST systems, and the results show that the proposed method is practical, effective and robust.

The obtained results are of practical significance for determining the optimal PM threshold and, thus, consists of the main contribution of this paper. The system aging property and the corresponding operation time models make the maintenance model more practical and more comfortable to implement, which is another contribution of this paper. In addition, the possible applications of the proposed method are much wider. For example, it can be used for other oil-lubricated systems with mission constraints, such as the wind turbine gearbox in wind power systems [23], and the diesel engine in marine freighters [11]. Moreover, the obtained outcomes can also complement the researches of system residual life prediction; for example, the derived results in the works of Liu et al. [25], Vališ et al. [19], Yan et al. [31]. Following the method proposed in this paper, these previous results might be complemented when used in lubricant condition monitoring, wear failure evaluation and other Prognostics and Health Management (PHM) applications.

The main contribution of this paper is not only a new direction in the maintenance policy optimization for oil-lubricated systems but also open up possibilities for evaluating the system residual damage using other historical operation data. There are several possible directions deserving future research. First, the mechanism-based system aging property description method might be developed to complement the historical failure data-based method proposed in this paper. Second, other distributions may have to be considered to establish the operating time model and the malignance time model in other cases. Third, more degradation modeling methods that can fuse multiple condition monitoring data may have to be used when modeling other systems. Fourth, other maintenance modeling method that can deal with the cases with storage conditions should be investigated in future research.

Acknowledgment

This work is partially supported by the NSFC under grant numbers 51475044, 51975047, and partially supported by the China Scholarship Council under grant number 201806030083.The authors have no conflicts of interest to declare.

References

1. Abbou A, Makis V. Group maintenance: A restless bandits approach. INFORMS Journal on Computing 2019; 31(4): 719-731, https://doi.org/10.1287/ijoc.2018.0863.
2. Ahn H S, Yoon E S, Sohn D G, Kwon O K, Shin K S, Nam C H. Practical contaminant analysis of lubricating oil in a steam turbine-generator. Tribology international 1996; 29(2): 161-168, https://doi.org/10.1016/0301-679X(95)00003-M.
3. Barde S R, Yacout S, Shin H. Optimal preventive maintenance policy based on reinforcement learning of a fleet of military trucks. Journal of Intelligent Manufacturing 2019; 30(1): 147-161, https://doi.org/10.1007/s10845-016-1237-7.
4. 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.
5. Du Y, Wu T, Makis V. Parameter estimation and remaining useful life prediction of lubricating oil with HMM. Wear 2017; 376: 1227-1233, https://doi.org/10.1016/j.wear.2016.11.047.
6. Du Y, Wu T, Zhou S, Makis V. Remaining useful life prediction of lubricating oil with dynamic principal component analysis and proportional hazards model. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology 2019; 1350650119874560, https://doi.org/10.1177/1350650119874560.
7. Fan C Y, Fan P S, Chang P C. A system dynamics modeling approach for a military weapon maintenance supply system. International Journal of Production Economics 2010; 128(2): 457-469, https://doi.org/10.1016/j.ijpe.2010.07.015.
8. Gajewski J, Vališ D. The determination of combustion engine condition and reliability using oil analysis by MLP and RBF neural networks. Tribology International 2017; 115: 557-572, https://doi.org/10.1016/j.triboint.2017.06.032.
9. Guinat E. Spectrometric oil analysis: atomic emission spectrometric analysis of wear metals in lube oils by RDE method. Tribology international 1985; 18(4): 246, https://doi.org/10.1016/0301-679X(85)90070-2.
10. Guo C, Wang W, Guo B, Si X. A maintenance optimization model for mission-oriented systems based on Wiener degradation. Reliability Engineering & System Safety 2013; 111: 183-194, https://doi.org/10.1016/j.ress.2012.10.015.
11. Jun H B, Kiritsis D, Gambera M, Xirouchakis P. Predictive algorithm to determine the suitable time to change automotive engine oil. Computers & Industrial Engineering 2006; 51(4): 671-683, https://doi.org/10.1016/j.cie.2006.06.017.
12. Li X, Makis V, Zuo H, Cai J. Optimal Bayesian control policy for gear shaft fault detection using hidden semi-Markov model. Computers & Industrial Engineering 2018; 119: 21-35, https://doi.org/10.1016/j.cie.2018.03.026.
13. Liao H, Elsayed E A, Chan L Y. Maintenance of continuously monitored degrading systems. European Journal of Operational Research 2006; 175(2): 821-835, https://doi.org/10.1016/j.ejor.2005.05.017.
14. Liu Y, Ma B, Zheng C S, Li S C. Research and evaluation on wear in power-shift steering transmission through oil spectral analysis with RKPCA method. Spectroscopy and Spectral Analysis 2015; 35(5): 1370-1375.
15. Liu Y, Ma B, Zheng C S, Xie S Y. Failure prediction of power-shift steering transmission based on oil spectral analysis with Wiener process. Spectroscopy and Spectral Analysis 2015; 35(9): 2620-2624.
16. Makis V, Wu J, Gao Y. An application of DPCA to oil data for CBM modeling. European Journal of Operational Research 2006; 174(1): 112-123, https://doi.org/10.1016/j.ejor.2005.03.010.
17. Myung I J. Tutorial on maximum likelihood estimation. Journal of Mathematical Psychology 2003: 47(1): 90-100, https://doi.org/10.1016/S0022-2496(02)00028-7.
18. Osman T, Velex P. Static and dynamic simulations of mild abrasive wear in wide-faced solid spur and helical gears. Mechanism and Machine Theory 2010; 45(6): 911-924, https://doi.org/10.1016/j.mechmachtheory.2010.01.003.
19. Prabhakaran A, Jagga C R. Condition monitoring of steam turbine-generator through contamination analysis of used lubricating oil. Tribology International 1999; 32(3): 145-152, https://doi.org/10.1016/S0301-679X(99)00028-6.
20. Raposo H, Farinha J T, Fonseca I, Galar D. Predicting condition based on oil analysis-A case study. Tribology International 2019; 135: 65-74, https://doi.org/10.1016/j.triboint.2019.01.041.
21. Scholz F W, Stephens M A. K-sample Anderson-Darling tests. Journal of the American Statistical Association 1987; 82(399): 918-924, https://doi.org/10.1080/01621459.1987.10478517.
22. Sheng S. Monitoring of wind turbine gearbox condition through oil and wear debris analysis: A full-scale testing perspective. Tribology Transactions 2016); 59(1): 149-162, https://doi.org/10.1080/10402004.2015.1055621.
23. Tian Z, Jin T, Wu B, Ding F. Condition based maintenance optimization for wind power generation systems under continuous monitoring. Renewable Energy 2011; 36(5): 1502-1509, https://doi.org/10.1016/j.renene.2010.10.028.
24. Vališ D, Gajewski J, Žák L. Potential for using the ANN-FIS meta-model approach to assess levels of particulate contamination in oil used in mechanical systems. Tribology International 2019; 135: 324-334, https://doi.org/10.1016/j.triboint.2019.03.012.
25. Vališ D, Žák L. Contribution to prediction of soft and hard failure occurrence in combustion engine using oil tribo data. Engineering Failure Analysis 2017; 82: 583-598, https://doi.org/10.1016/j.engfailanal.2017.04.018.
26. Vališ D, Žák L, Pokora O. Perspective approach in using anti-oxidation and anti-wear particles from oil to estimate residual technical life of a system. Tribology International 2018; 118: 46-59, https://doi.org/10.1016/j.triboint.2017.09.017.
27. Vališ D, Žák L, Pokora O, Lánský P. Perspective analysis outcomes of selected tribodiagnostic data used as input for condition based maintenance. Reliability Engineering & System Safety 2016; 145: 231-242, https://doi.org/10.1016/j.ress.2015.07.026.
28. Wakiru J M, Pintelon L, Muchiri P N, Chemweno P K. A review on lubricant condition monitoring information analysis for maintenance decision support. Mechanical Systems and Signal Processing 2019; 118: 108-132, https://doi.org/10.1016/j.ymssp.2018.08.039.
29. Wang W. A prognosis model for wear prediction based on oil-based monitoring. Journal of the Operational Research Society 2007; 58(7): 887-893, https://doi.org/10.1057/palgrave.jors.2602185.
30. Wang W, Hussin B, Jefferis T. A case study of condition based maintenance modelling based upon the oil analysis data of marine diesel engines using stochastic filtering. International Journal of Production Economics 2012; 136(1): 84-92, https://doi.org/10.1016/j.ijpe.2011.09.016.
31. Yan S, Ma B, Zheng C. Health index extracting methodology for degradation modelling and prognosis of mechanical transmissions. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 137-144, https://doi.org/10.17531/ein.2019.1.15.

32. Yan S. F., Ma B., Zheng C. S. Condition-based maintenance for power-shift steering transmission based on oil spectral analysis. Spectroscopy and Spectral Analysis 2019; 39(11): 3470-3474.
33. Yan S, Ma B, Zheng C, Chen J. An optimal lubrication oil replacement method based on selected oil field data. IEEE Access 2019; 7: 92110-92118, https://doi.org/10.1109/ACCESS.2019.2927426.
34. Yan S F, Ma B, Zheng C S, Zhu L A, Chen J W, Li H. Z. Remaining useful life prediction of power-shift steering transmission based on uncertain oil spectral data. Spectroscopy and Spectral Analysis 2019; 39(2): 553-558.
35. Zheng C, Liu P, Liu Y, Zhang Z. Oil-based maintenance interval optimization for power-shift steering transmission. Advances in Mechanical Engineering 2018; 10(2): 1687814018760921, https://doi.org/10.1177/1687814018760921.
36. Zhu J, Yoon J M, He D, Bechhoefer E. Online particle-contaminated lubrication oil condition monitoring and remaining useful life prediction for wind turbines. Wind Energy 2015; 18(6): 1131-1149, https://doi.org/10.1002/we.1746.

Shufa YAN
Biao MA
Xu WANG
Jianhua CHEN
Changsong ZHENG
School of Mechanical Engineering
Beijing Institute of Technology
5 South Zhongguancun Street, Haidian District, Beijing 100081, China

E-mails: 3120160206@bit.edu.cn, 3220190331@bit.edu.cn,
1324471347@qq.com, 3220185033@bit.edu.cn,
zhengchangsong@bit.edu.cn

Article citation info:
KOWALCZYK J, MADEJ M, OZIMINA D. Evaluation of performance characteristics of the environmentally friendly cutting fluid with zinc aspartate. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 465–471, http://dx.doi.org/10.17531/ein.2020.3.9.

Joanna KOWALCZYK
Monika MADEJ
Dariusz OZIMINA

EVALUATION OF PERFORMANCE CHARACTERISTICS OF THE ENVIRONMENTALLY FRIENDLY CUTTING FLUID WITH ZINC ASPARTATE

OCENA WŁAŚCIWOŚCI EKSPLOATACYJNYCH PROEKOLOGICZNEJ CIECZY CHŁODZĄCO-SMARUJĄCEJ ZAWIERAJĄCEJ ASPARAGINIAN CYNKU*

The effect of the cutting fluid with zinc aspartate on the quality of the workpiece surface layer is reported. Until now, zinc aspartate has been used primarily in medicine and pharmacology. This paper compares the ecological cutting fluid containing zinc aspartate with a classic mineral oil-based coolant. Toxicity tests and a controlled process of tool wear during face turning were performed. Test results indicate that the use of zinc aspartate-based cutting fluids contributes to the reduction of the material roughness parameter values up to 35%, benefitting the final quality of the workpiece.

***Keywords**: ecological cutting fluid, zinc aspartate, tool wear, surface topography, microbiological tests.*

W pracy przedstawiono wyniki badań wpływu cieczy chłodząco-smarującej z asparaginianem cynku na jakość technologiczną warstwy wierzchniej obrabianych elementów. Asparaginian cynku dotychczas nie był stosowany w takich rozwiązaniach, głównie wykorzystywany był w medycynie i farmakologii. W badaniach przeprowadzono analizę porównawczą proekologicznego chłodziwa zawierającego asparaginian cynku z klasycznym chłodziwem opartym na bazie oleju mineralnego. Ciecze chłodząco-smarujące poddano badaniom toksyczności oraz wykonano kontrolowany proces eksploatacji narzędzi w czasie toczenia poprzecznego. Wyniki badań wskazują, że zastosowanie chłodzenia cieczą na bazie asparaginianu cynku redukuje parametry chropowatości obrabianego materiału nawet o 35%, korzystnie wpływając na jakość finalną detalu.

***Słowa kluczowe**: proekologiczna ciecz chłodząco-smarująca, asparaginian cynku, zużycie narzędzia, topografia powierzchni, badania mikrobiologiczne.*

1. Introduction

Machining is the most commonly used method of manufacturing components and parts in virtually all production technologies [18, 20]. Machining operations use cutting fluids [2, 23] to cool, lubricate and improve the surface condition of the machined part [22], to transport chips from the machining zone [3] and temporarily protect the product against corrosion [6, 21]. In order to fulfil their tasks, cutting fluids must have a special chemical composition, which cannot adversely affect people or the environment [2, 6, 24]. The most widely used cutting fluids are based on water or oil. Despite the obvious advantages, including low cost, mineral oil-based fluids are toxic for the environment and their degradation processes are difficult to control [18, 20]. In addition, friction modifiers, extreme pressure (EP) and anti-wear (AW) additives, corrosion inhibitors, antioxidants, etc. [1, 3, 10, 16] used in the fluids may add to their harmful effects.

Yadav et al. [21] examined the impact of EP and AW additives in new and used engine oils class SAE 15W40 and SAE 20W50 on the wear of the test balls. The tests were carried out on a Ducom TR-30L four-ball tester at a temperature of 75ºC, under a load of 392 N and at a constant speed of 1250 rpm. The results indicated that the wear scare diameter after tests increased gradually with increasing oil use time. It was found that the effect of EP and AW additives was primarily determined by the engine oil operating conditions.

Maruda et al. [9] studied the impact of EP and AW additives on the topography of steel surfaces during MQCL turning with various coolants at varying flow rates. The results showed that the addition of phosphate ester-based additives to the active medium resulted in the formation of a tribofilm at the tool-chip interface, which reduced friction. The lubrication method using a minimum amount of coolant with EP/AW substances improved surface topography parameters.

In another article, Maruda et al. [10] compared dry machining and cooling with compressed air, emulsion mist, emulsion mist + Crodafos O4A-LQ-(MH) and emulsion mist + Crodafos EHA-LQ-(MH). They found that cooling with emulsion mist and modifiers can provide an 80% reduction of roughness parameters.

Maruda et al. [8] reported the results from introducing the additives in the emulsion mist first on the processed surface and then to the contact zone of the friction pair. The quality of the machined stainless steel surface improved as a result of reduced roughness, and phosphate ester-based additives used in the emulsion mist intensified this effect. The phosphate-based additives remained on the machined surface even after 30 minutes of machining under high loads, temporarily lowering the friction coefficient and temperature in the interfacial zone.

Lubricants based on vegetable oil are introduced to the market for their biodegradability and non-toxicity. They are a potential lubricant preparation alternative that does not use petroleum derivatives [15, 17]. Much work has been devoted to research on cutting fluids based on vegetable oils.

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

Ozcelik et al. [14] tested coolants based on sunflower and rapeseed oil with an EP additive at 8% and 12%. The metalworking fluid containing rapeseed oil provided a better surface quality of the workpiece than that based on sunflower oil by more than 11% and 4%, respectively. Kumar et al. [5] studied coconut oil coolant with EP additives. They showed that coconut oil reduced the feed force by 31%, the pressure force by 28%, the cutting force by 20%, the cutting tool temperature by 7% and the wear of the tool face by 34% compared to other machining fluids.

In turn, Zhang et al. [24, 25] tested soybean coolants, which were compared to petroleum-based machining fluids and dry machining. They showed that the soy-based coolant worked similarly to the petroleum-based fluid. At the same time, they obtained much better surface roughness and less tool wear, compared with coolant-free machining.

Trajano et al. [19] performed tribological tests of sunflower and soybean oils containing CuO additives and ZnO nanoparticles. In the case of soybean nanofluids with CuO and ZnO concentrations of 0.5 wt.%, the friction coefficient decreased by 11% and 18% respectively. However, for the sunflower oil with the same filler fraction, the coefficient of friction decreased for CuO and ZnO by 22% and 20%, respectively.

The literature review above shows that there are large gaps in coolant research. To date, no studies have been conducted to determine the effect of zinc aspartate as a biodegradable modifying additive for cutting fluids. Bibliometric analysis indicates the use of zinc aspartate only in medicine and pharmacology. The use of zinc aspartate as an additive to cutting fluids should improve the lubricating properties and maintain the machine and equipment in an appropriate operating condition. [4].

This paper presents the results obtained during the machining operation with the use of a non-toxic cutting fluid containing zinc aspartate. Its impact on selected factors and surface quality of the machined workpiece was also determined.

2. Materials

2.1. Metalworking fluids

The fluid used in the tests was an eco-friendly, demineralized water (DEMI) based cutting fluid with 5% zinc aspartate and, among others:

- alkanolamine borate;
- a biodegradable oligomer based on poly(aspartic acid) (PASA); and
- demineralized water.

Zinc polyaspartates ensure the fluid biostability. Physical and chemical properties of the cutting fluid used are summarized in Table 1.

A zinc aspartate-containing coolant was compared to a commercial cutting fluid based on mineral oil and containing ethoxylated aliphatic alcohols with chain lengths C16-18, boric acid, dicyclohexylamine, 3-iodo-2-propynyl butylcarbamate, and 1,2-benzisothiazol-3 (2H)-one. It is used for general to heavy machining of aluminium, steel, cast iron, non-ferrous metals, aluminium alloys, brass and copper. The fluid contributes to the good quality of machined surfaces. The basic parameters of the fluid are compiled in Table 2.

2.2. Machined materials and tools

A tool made of HS6-5-2C steel was used for turning the front faces on the workpiece. This steel is characterized by very good ductility, impact strength and abrasion resistance. It is designed for hot work and can be subjected to heat treatment at elevated temperatures: hardening - 1190 ÷ 1230 ° C and tempering - 550 ÷ 650 ° C. Its hardness after heat treatment at 500 ÷ 550 °C is 65 HRC. The chemical composition of HS6-5-2C steel is shown in Table 3.

The HS6-5-2C steel tools were chosen for their common use in machining (about 40% of all tools used in manufacturing) Despite the growing demand for cemented carbide tools that can be used at high speeds on CNC machines, high speed tool steel is still an economical alternative, given the price and easy sharpening. Steels obtained by sintering high-speed steel powders are increasingly used for tools, whose blades are coated with complex coatings dedicated to specific applications.

The workpiece was a 38 mm diameter roller made of C45 steel, which is a non-alloy, medium carbon steel for quenching and tempering, difficult to weld, easy to machine, with the chemical composition as in Table 4. Products made of C45 can be surface hardened to 50 ÷ 60 HRC.

3. Methods

A FTIR Spectrum Two infrared spectrometer with the Perkin Elmer ATR adapter was used to test the thermo-oxidative kinetics of operating fluids. Oligomer samples were measured as pure and after model tribological tests. The following analytical parameters were used during spectral tests:

Table 1. Properties of the zinc aspartate-based cutting fluid

Colour	Odour	*pH*, 3%	Density, g/cm^3	Water solubility
orange to red	specific	9.2 ÷ 9.7	1.20 ÷ 1.25	soluble

Table 2. Main parameters of the cutting fluid based on mineral oil

Colour	Odour	Mineral oil content	*pH*, 5%	Density, g/cm^3	Water solubility
yellow-brown	that of mineral oil	56%	9.1	0.92 ÷ 0.96	soluble

Table 3. Composition of HS6-5-2C steel

Chem. element	C	Mn	Si	P	S	Cr	Ni	Mo	W	V	Co	Cu
Content, %	0.82÷ 0.92	max 0.4	max 0.5	max 0.03	max 0.03	3.5÷ 4.5	max 0.4	4.5÷ 5.5	6 ÷ 7	1.7÷ 2.1	max 0.5	max 0.3

Table 4. Chemical composition of C45 steel

Chem. element	C	Mn	Si	P	S	Cu	Cr	Ni	Mo	Cu
Content, %	0.42÷ 0.5	0.5÷ 0.8	0.1÷ 0.4	max 0.04	max 0.04	max 0.3	max 0.3	max 0.3	max 0.1	max 0.3

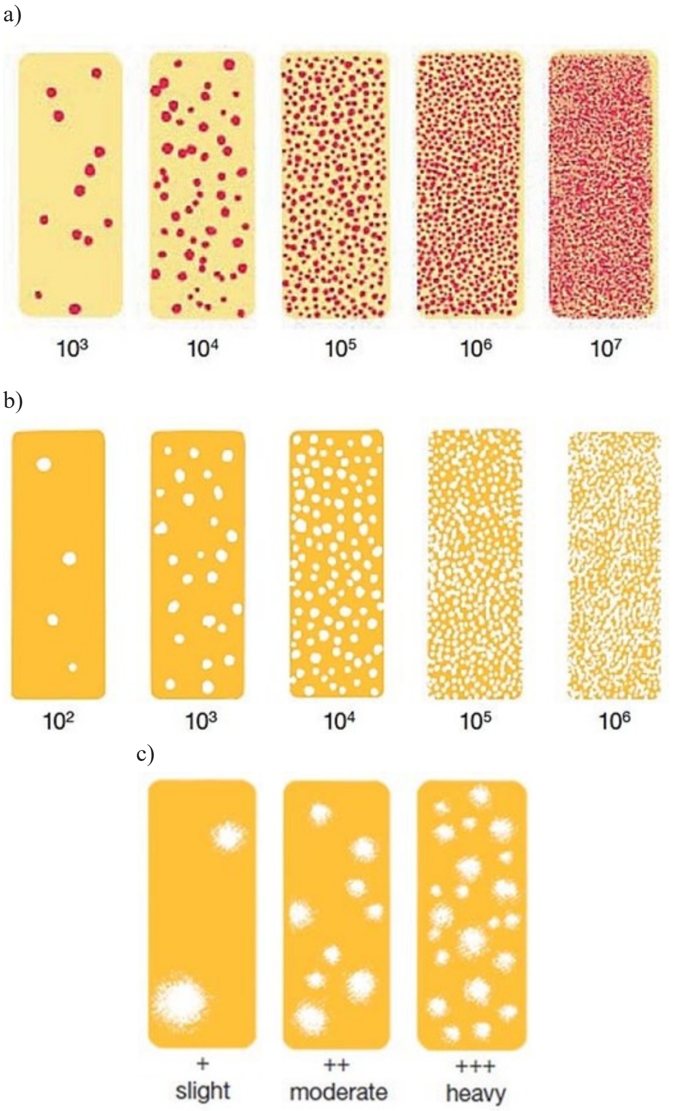


Fig. 1. Patterns of colonies of microorganisms: a) bacteria, b) yeast, c) mould [12]

- spectral range: 4000 ÷ 400 cm^{-1};
- number of background scans: air and samples – 4.

For the microbiological assessment of fluid toxicity, a special Microbiology Cult Dip Combi kit was used to determine the presence of microorganisms. It consisted of containers with attached test plates covered with appropriate media. One side of the test plates was lighter and was used to detect the presence of bacteria, and the other side was darker to detect the presence of yeast and fungi. All parts of the test kit were sterile. Toxicity tests according as per instructions [12] were applied to the zinc aspartate- containing fluid and the conventional mineral oil-based coolant. The samples were checked first after 48 hours, then after 96 hours and finally after 7 days. At 7 days the samples were subjected to organoleptic assessment and compared with microorganism colony standards (Fig. 1).

Table 5. Parameters of the cutting process

Cutting speed, v_c, m/min	Feed per revolution, *f*, mm/rev	Cutting depth, a_p, mm
47.5 ÷ 0	0.098	0.5

Facing was performed on a CTX 310 ECO numerically controlled lathe by DMG MORI using the Sinumerik 810 control system. By maintaining a constant rate of rotation at 400 m/min at each pass, the cutting speed v_c changed cyclically in the range 47.5 ÷ 0 m/min. In order to compare the properties of the fluid with zinc aspartate to those of the mineral oil-based coolant, face turning was performed using both fluids. After the operation, tool wear and built-up edge were measured. Turning parameters are shown in Table 5.

Replaceable cutting tools were used - insert tool bits with a square cross-section 10 mm x 10 mm, held in a bit socket. The inserts were made of high speed steel HS6-5-2C. The geometry parameters were as follows:

- major cutting edge angle $K_r = 36.6°$ in the reference plane P_r,
- minor cutting edge angle $K_r' = 53.4°$,
- positive rake angle = 5.3°,
- tool included angle $\varepsilon_r = 90°$,
- cutting edge inclination angle $\lambda_s = 7°$ in the cutting edge plane P_s,
- clearance angle $\alpha = 5.3°$,
- rake angle $\gamma = 7°$,
- wedge angle $\beta = 77.7°$,
- nose radius $r_s = 0.04$ mm.

The workpiece material was a C45 steel roller with a diameter of 38 mm. During turning, conventional flood cooling was applied to the tool rake face. The view and diagram of the coolant supply and discharge system is shown in Figure 2.

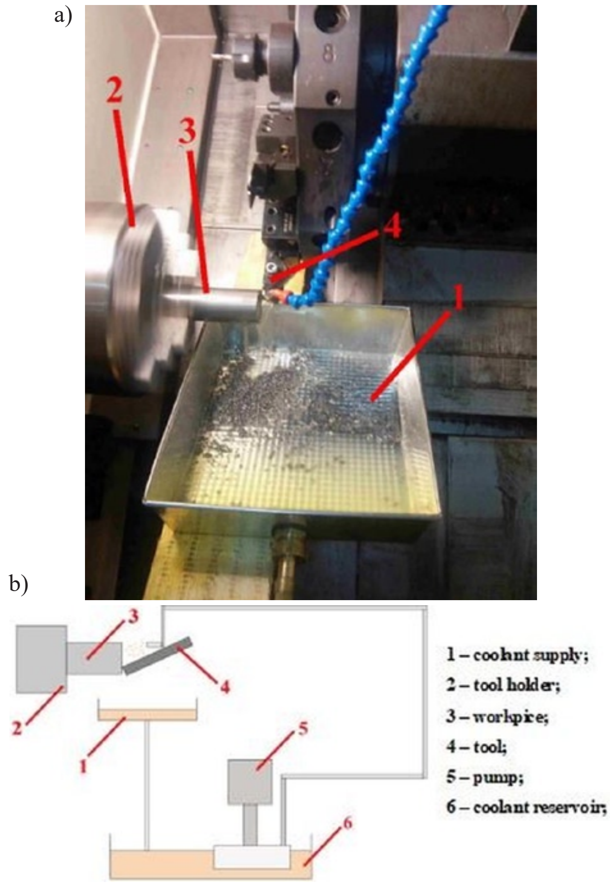


Fig. 2. External coolant circulation system during face turning on a CTX 310 ECO lathe: a) view, b) diagram

Ten stages of face turning were performed. The first stage consisted of 10 passes, and 10 more passes were added in each subsequent stage (Fig. 3). After each stage, a thin piece – a "slice" of the workpiece - was cut off and the cutting tool was replaced.

The aim of the study was to assess the basic operating characteristics of the new, zinc aspartate-based cutting fluid and the wear of the cutter during face turning.

A JEOL JSM-7100F scanning electron microscope was used to examine the surface topography of workpieces and cutting tools. The EDS microanalyzer enabled the identification of chemical elements on the surface of cutting tools at the built-up edge after the 10th turning cycle with coolants.

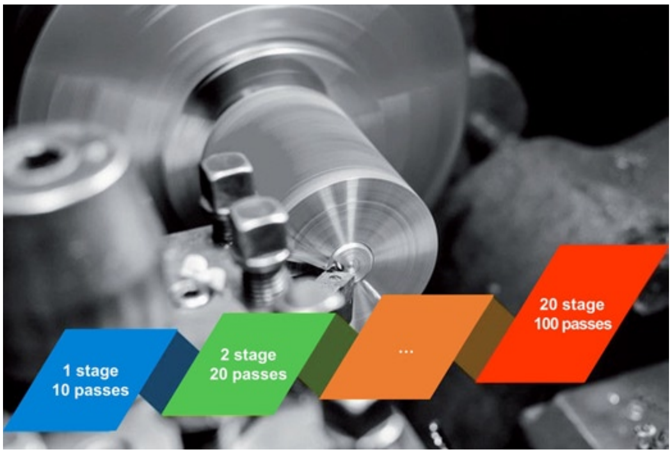


Fig. 3. Diagram of the face turning process for cutting tools

Geometry of the workpiece after turning was imaged using a DCM8 confocal microscope from Leica. In addition, the SX80 stereoscopic inspection microscope was used to observe the wear of the cutting tools after machining.

4. Results

Figure 4 compiles FTIR spectra which show a clear intensification of two bands with increasing load and friction path. The low intensity signals with a peak maximum at about 1394 and 1066 cm^{-1} correspond to C-O (COH) stretching vibrations and indicate the initiation of coolant degradation processes and the formation of a carboxyl group. Observations of the cutting fluid containing zinc aspartate after tribological tests reveal a change in its colour. With increasing load and friction path, the colour becomes darker, from light yellow to brown.

The cutting fluid can be regenerated by supplementing the basic component - zinc aspartate and controlling other operating parameters.

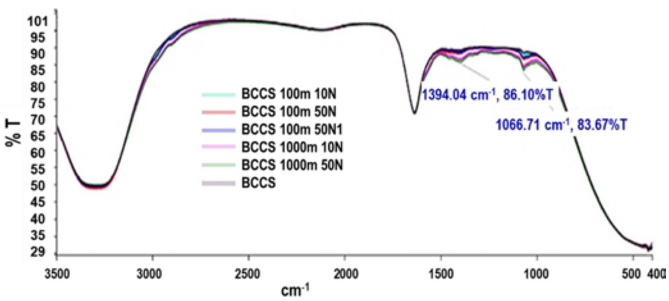


Fig. 4. FTIR spectra for the cutting fluid before and after tests

Toxicity of the zinc aspartate-based coolant and, for comparison, of the mineral oil-based fluid was evaluated using the Microbiology Cult Dip Combi kit. Two samples of each fluid were used in the tests. Observations began after 7 days (Figs. 5 and 6).

Fig. 5. View of test samples after 7 days – bacteria for the cutting fluid: a) containing zinc aspartate, b) containing mineral oil

Fig. 6. View of test samples after 7 days – yeast and molds for the cutting fluid: a) containing zinc aspartate, b) containing mineral oil

The cutting fluid containing zinc aspartate was found to be non-toxic. No colonies of bacteria (Fig. 5a), yeast or mould (Fig. 6a) were formed. However, on one of the slides earlier immersed in the mineral oil-based coolant several spots on the bacterial count side were observed (Fig. 5b). According to the guidelines [12], these spots indicate a minor infection. Considering the results obtained, the zinc aspartate coolant turned out to be better than the mineral oil-based coolant.

Using the inspection microscope software, wear measurements on the flank face were carried out in accordance with the standard [13], as shown in Fig. 7.

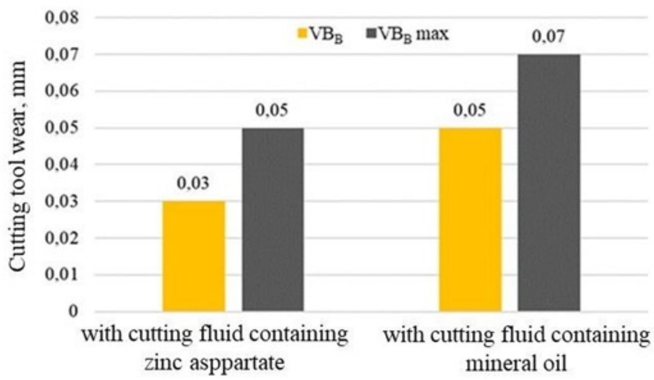


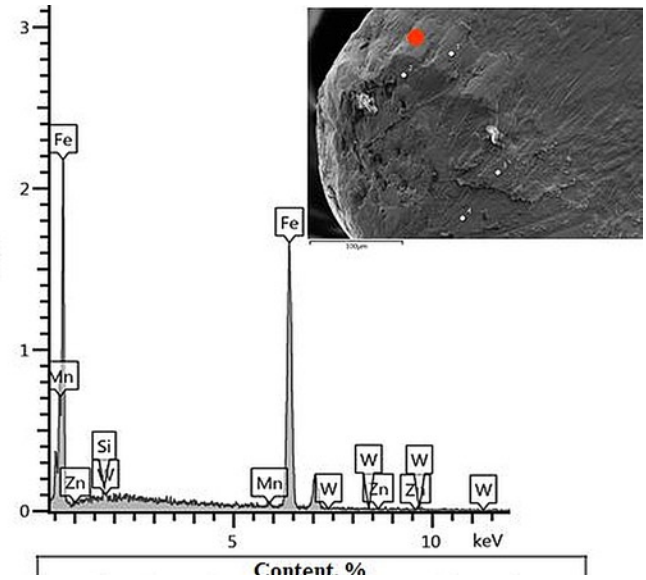
Fig. 7. Wear indices of cutting tools made of HS6-5-2C steel after turning with cutting fluids

The charts above show that the average VB_B and maximum VB_B*-max* values of the wear bandwidth were 0.03 mm and 0.05 mm respectively after the treatment with the coolant containing zinc aspartate. The tool wear parameters after turning with the coolant containing mineral oil were higher.

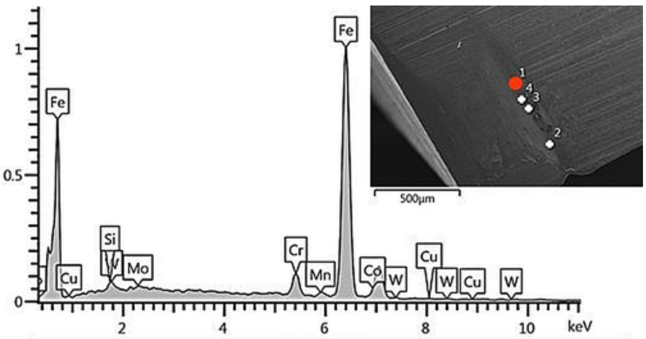
Figure 8 presents SEM images of cutting blades wear after dry machining and machining with the use of coolants, and the results of chemical composition analysed in micro-areas using the EDS method. After the final 10th turning test with coolants, built-ups formed in the place of wear on the cutting tools. Chemical composition of this excessive material was examined by scanning electron microscopy.

After turning with the zinc aspartate containing cutting fluid, a concentration of zinc atoms was recorded in addition to the elements included in the tool. This indicates that a thin surface layer of zinc compounds with anti-wear properties was formed as a result of tribochemical processes that occur mainly between improvers, contained mainly in lubricants, and friction surfaces. The speed and type of chemical reactions depend on the operating conditions of the friction node. The surface reaction layer produced in this way changes the working conditions of the friction pairs, which in turn leads to further tribochemical reactions. This surface layer reduces friction and extends the service life of friction pairs [16].

a)



b)

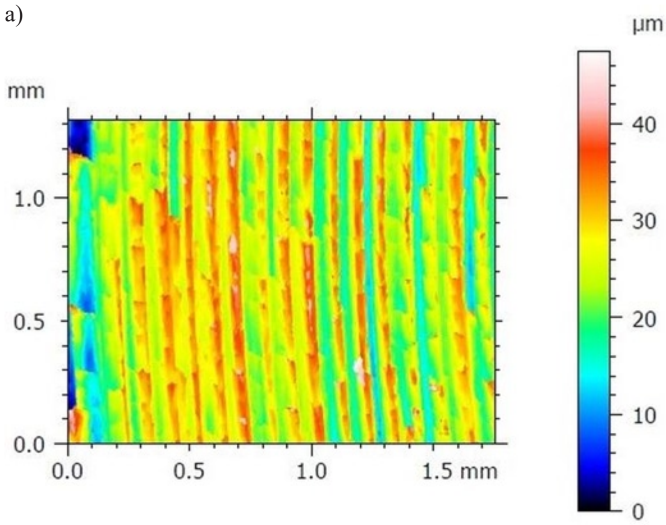


Content, %										
Si	V	Cr	Mn	Fe	Co	Ni	Cu	Mo	W	Total
0.76	-	4.34	0.82	89.12	0.83	-	0.80	0.49	0.65	97.37

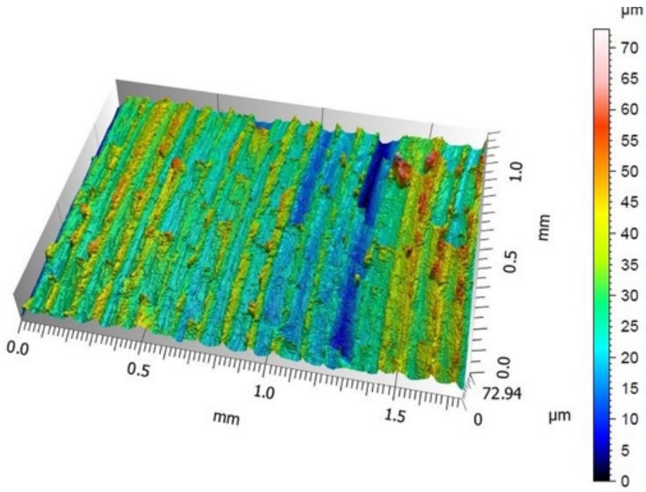
Fig. 8. X-ray analysis of wear area of the cutting tool edge after turning with the cutting fluid containing: a) zinc aspartate, b) mineral oil

Figures 9 and 10 show the topography and surface roughness profiles of workpieces in selected areas, obtained after stage 10 of dry turning and that with the cutting fluids.

a)



b)



c)

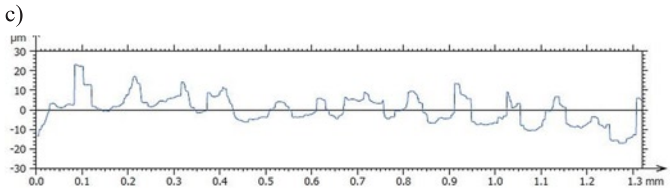


Fig. 9. Geometric structure of the workpiece edge after turning with the cutting fluid containing mineral oil: a) contour map, b) isometric image, c) primary profile

A comparison of the contour maps, isometric views and primary profiles reveals an even distribution of vertices [7] almost every 1 mm (feed: 0.098 mm) in both cases. The lowest elevations of approx. 16 µm and the shallow cavities of approx. 15 µm were formed on the workpiece after turning with the fluid containing zinc aspartate. In contrast, the highest elevations/peaks and depressions/valleys of about 20 µm were recorded after turning with the coolant containing mineral oil.

Table 6 shows the parameters of the geometric structure of machined surfaces formed after turning with the use of zinc aspartate and mineral oil cutting fluids.

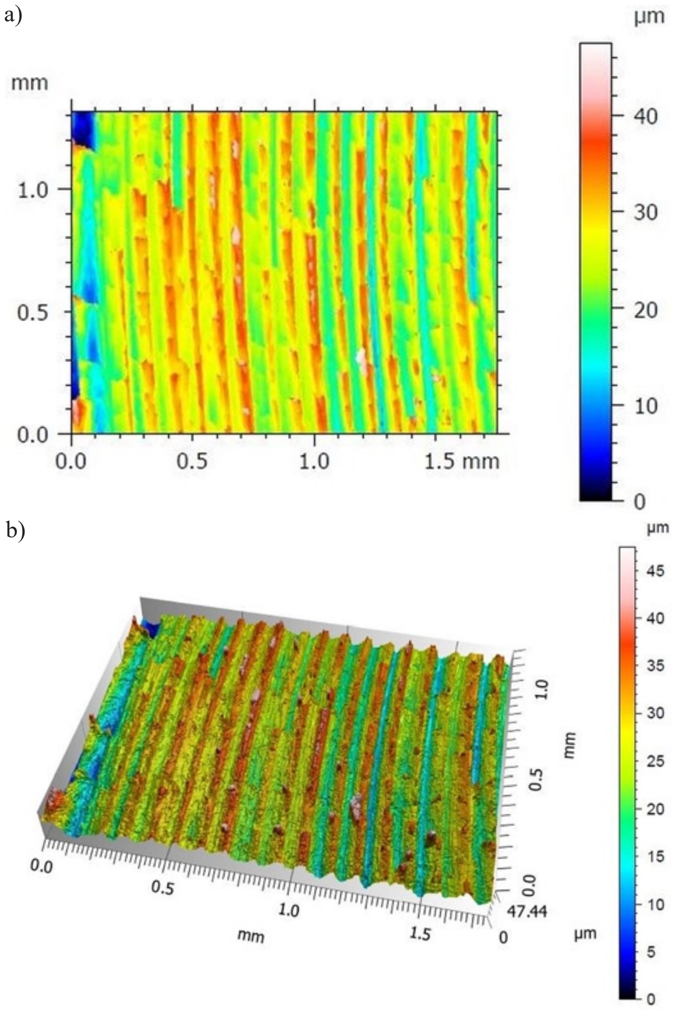
The values of surface roughness parameters of machined elements after the turning process with zinc aspartate containing coolant are lower than those for the mineral oil coolant. This indicates that the cutting fluid improves the surface quality of the workpiece.

4. Conclusion

The conclusions based on the test results are the following:

1. The cutting fluid with zinc aspartate used in the turning process reduced the wear of elements of the tribological system by forming layers containing zinc compounds with good anti-wear properties.
2. The workpiece subjected to turning with the zinc aspartate containing fluid had the lowest elevations and shallowest recesses after the process.
3. Surface morphology in the turning process depends on the direction of movement and size of the feed.
4. The non-toxic coolant with zinc aspartate was undergoing transformation as a result of the reactions dependent on the operating conditions.
5. The cutting fluid containing zinc aspartate does not show any toxic properties and can be subjected to simple regeneration.
6. Due to its good ecological properties, the tested cutting fluid with zinc aspartate can be used during operation on conventional lathes, where the operator's direct contact with the coolant is greater than on numerically controlled lathes.
7. The cutting fluid with zinc aspartate provided better technological quality of the surface layer of machined elements and extended the life of high speed steel tools compared to classic coolants. The surface roughness values after the use of zinc asparatate cutting fluid were lower (by 22% for Sa, by 25% for Sq, by half for Sp, by 11% for Sv, by 35% for Sz, more than 2 times for Ssk, and by 17% for Sku).

a)



b)

c)

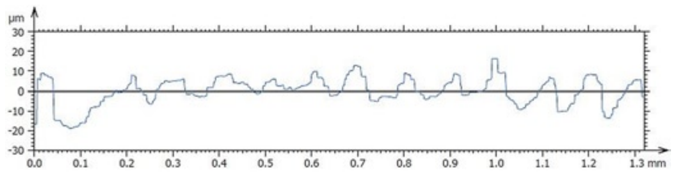


Fig. 10. Geometric structure of the workpiece edge after turning with a cutting fluid containing zinc aspartate: a) contour map, b) isometric image, c) primary profile

Table 6. Surface texture parameters obtained for the discs after turning

Turning conditions	Surface texture parameters						
	Sa	Sq	Sp	Sv	Sz	Ssk	Sku
	µm	µm	µm	µm	µm	-	-
with the mineral oil-based coolant	6.41	8.44	44.32	28.62	72.94	0.26	4.06
with the zinc aspartate-based coolant	5.00	6.31	22.12	25.32	47.44	-0.22	3.34

References

1. Brinksmeier E, Meyer D, Huesmann-Cordes A G, Herrmann C. Metalworking fluids - Mechanisms and performance. CIRP Annals - Manufacturing Technology 2015; 64(2): 605-628, https://doi.org/10.1016/j.cirp.2015.05.003.
2. Jain A, Kumar S, Bajpai V, Wook Park H. Replacement of Hazard Lubricants by Green Coolant in Machining of Ti6Al4V: A 3D FEM Approach. The International Journal of Precision Engineering and Manufacturing 2019; 20 (6): 1027-1035, https://doi.org/10.1007/s12541-019-00111-2.
3. Kępczak N, Rosik R, Pawłowski W, Sikora M, Witkowski B, Bechciński G, Stachurski W. The dynamics of wear of cutting inserts during turning of non-homogeneous material on the example of polymer concrete. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 3 (20): 478-483, https://doi.org/10.17531/ein.2018.3.18.
4. Kowalczyk J. Ocena właściwości wybranych proekologicznych cieczy chłodząco-smarujących w systemach tribologicznych z powłokami przeciwzużyciowymi. Rozprawa doktorska, Kielce: PŚk Pubisher, 2019.
5. Kumar B S, Padmanabhan G, Krishna P V. Experimental Investigations of Vegetable Oil Based Cutting Fluids with Extreme Pressure Additive in Machining of AISI 1040 Steel. Manufacturing Science and Technology 2015, 3: 1-9.
6. Li K M, Liang S Y. Performance profiling of minimum quantity lubrication in machining. The International Journal of Advanced Manufacturing

Technology 2007; 35: 226-233, https://doi.org/10.1007/s00170-006-0713-1.
7. Maruda R, On the surface roughness of 316L austenitic stainless steel after turning when cooling with MQCL method. Mechanik 2016; 8-9: 1058-1059, https://doi.org/10.17814/mechanik.2016.8-9.246.
8. Maruda R W, Feldshtein E E, Legutko S, Królczyk G M. Improving the efficiency of running-in for a bronze-stainless steel friction pair. Journal of Friction and Wear 2015; 36 (6): 548-553, https://doi.org/10.3103/S1068366615060082.
9. Maruda R W, Królczyk G M, Wojciechowski Sz, Żak K, Habrat W, Niesłony P. Effects of extreme pressure and anti-wear additives on surface topography and tool wear during MQCL turning of AISI 1045 steel. Journal of Mechanical Science and Technology 2018; 32 (2): 1585-1591, https://doi.org/10.1007/s12206-018-0313-7.
10. Maruda R W, Legutko S, Królczyk G M, Hloch S, Michalski M. An influence of active additives on the formation of selected indicators of the condition of the X10CrNi18-8 stainless steel surface layer in MQCL conditions. International Journal of Surface Science and Engineering 2015; 9 (5): 452-465, https://doi.org/10.1504/IJSURFSE.2015.072069.
11. McNutt J, He Q. Development of biolubricants from vegetable oils via chemical modification. Journal of Industrial Engineering Chemistry 2016; 36: 1-12, https://doi.org/10.1016/j.jiec.2016.02.008.
12. Microbiology - 100778 Cult-Dip combi - instrukcja badania toksyczności, test bakteriologiczny.
13. Norma PN-ISO 3685 Badanie trwałości noży tokarskich punktowych.
14. Ozcelik B, Kuram E, Huseyin C M, Demirbas E. Experimental investigations of vegetable based cutting fluids with extreme pressure during turning of AISI 304L. Tribology International 2011, 44: 1864-1871, https://doi.org/10.1016/j.triboint.2011.07.012.
15. Park K H, Yang G D, Lee D Y. Tool wear analysis on coated and uncoated carbide tools in inconel machining. International Journal of Precision Engineering and Manufacturing 2015; 16(7): 1639-1645, https://doi.org/10.1007/s12541-015-0215-x.
16. Płaza S, Margielewski L, Celichowski G. Wstęp do tribologii i tribochemia, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2005.
17. Sartori S, Taccin M, Pavese G, Ghiotti A, Bruschi S. Wear mechanisms of uncoated and coated carbide tools when machining Ti6Al4V using LN2 and cooled N2. The International Journal of Advanced Manufacturing Technology 2018; 95(1 - 4): 1255-1264, https://doi.org/10.1007/s00170-017-1289-7.
18. Styp-Rekowski M, Matuszewski M, Oborski I L. Possibilities of some constructional materials cutting by means of water-abrasive jet. Journal of Polish CIMAC 2012; 7 (3): 293-300.
19. Trajano M F, Moura E I F, Ribeiro K S B, Alves S M. Study of oxide nanoparticles as additives for vegetable lubricants. Journal of Materials Research 2014, 17, 1124-1128, https://doi.org/10.1590/1516-1439.228213.
20. Wartacz A, Świć A, Zubrzycki J. The waste of tool blade, and a change of turning parameters in next cuts. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2004; 4 (24): 58-61.
21. Yadav G, Tiwari S, Jain M L. Tribological analysis of extreme pressure and anti-wear properties of engine lubricating oil using four ball tester. Materials Today: Proceedings 2018; 5: 248-253, https://doi.org/10.1016/j.matpr.2017.11.079.
22. Yoon H S, Kim M S, Jang K H, Ahn S H. Future perspectives of sustainable manufacturing and applications based on research databases. The International Journal of Precision Engineering and Manufacturing 2016; 17(9): 1249-1263, https://doi.org/10.1007/s12541-016-0150-5.
23. Zaleski K. The metod of investigations of lubricating properties of cutting fluids, Eksploatacja i Niezawodnosc - Maintenance and Reliability 2004; 4 (24): 62-65.
24. Zhang J L, Rao P N. Green/Sustainable Manufacturing-Evaluation of a Soybean-Based Metal Cutting Fluid in Turning Operation. Applied Mechanics and Materials 2013, 392: 925-930, https://doi.org/10.4028/www.scientific.net/AMM.392.925.
25. Zhang J, Rao P N, Eckman M. Evaluation of bio-based cutting fluids in using multiple quality characteristics through Taguchi design method. International Journal Modern Engineering 2012, 12: 35-44.

Joanna KOWALCZYK
Monika MADEJ
Dariusz OZIMINA
Department of Mechanical Design
Kielce University of Technology,
Al. 1000-lecia Państwa Polskiego 7, 25-314 Kielce, Poland

E-mails:jkowalczyk@tu.kielce.pl, mmadej@tu.kielce.pl, ozimina@tu.kielce.pl

Article citation info:
ANDREJIOVA M, GRINCOVA A, MARASOVA D. Analysis of tensile properties of worn fabric conveyor belts with renovated cover and with the different carcass type. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 472–481, http://dx.doi.org/10.17531/ein.2020.3.10.

Miriam ANDREJIOVA
Anna GRINCOVA
Daniela MARASOVA

ANALYSIS OF TENSILE PROPERTIES OF WORN FABRIC CONVEYOR BELTS WITH RENOVATED COVER AND WITH THE DIFFERENT CARCASS TYPE

ANALIZA WŁAŚCIWOŚCI WYTRZYMAŁOŚCIOWYCH ZUŻYTYCH TKANINOWYCH TAŚM PRZENOŚNIKOWYCH Z RÓŻNYMI TYPAMI RDZENIA PO RENOWACJI GÓRNEJ OKŁADKI

Conveyors are the means of transportation used in many industries. The load-bearing and tractive component of a belt conveyor is a conveyor belt which consists of a carcass and cover layers. During an operation, belts are exposed to loads that cause damage to the belts. It is therefore necessary to ensure that a conveyor belt possesses required mechanical properties during the transport of material. The key mechanical properties of a conveyor belt are tensile properties. They are significantly affected by the fabric carcass of a conveyor belt. The tensile properties of conveyor belts are largely affected by the carcass materials. They are also affected by the types of fibres in the longitudinal (warp) and transverse (weft) directions of the fabric carcass because the carcass transfers all tensile stresses of the conveyor belt. A fabric conveyor belt is regarded as a composite material, consisting of the carcass (polyamide P and polyamide-polyester EP) and the cover layers. The costs of a conveyor belt represent 10-30 % of the price of the entire conveyor. It is therefore reasonable to prefer only those conveyor belts that show the properties prescribed by relevant norms. The subject of the article is worn conveyor belts with renovated top cover (renovated conveyor belts). The tensile properties are used to assess the suitability for further use of renovated conveyor belts in practice. The article presents the analysis of tensile properties of renovated fabric conveyor belts in relation to the carcass type. The observed results were compared applying the DOE method, regression and correlation analysis, and the method of statistic induction. All the conclusions, made based on the above-listed methods, are identical for the examined tensile properties. The results indicate that the examined tensile properties of conveyor belts have not undergone any significant change after the renovation of the top cover layer.

Keywords: *tensile properties, worn conveyor belt, renovation of top cover, carcass type.*

Przenośniki taśmowe znajdują zastosowanie jako urządzenia transportowe w wielu gałęziach przemysłu. Nośnym i pociągowym elementem przenośnika taśmowego jest taśma transportowa, która zbudowana jest z rdzenia i okładek. Podczas pracy, taśmy narażone są na obciążenia, które prowadzą do ich uszkodzenia. Dlatego konieczne jest aby taśma transportująca materiał posiadała wymagane właściwości mechaniczne. Kluczowymi właściwościami mechanicznymi taśmy przenośnikowej są właściwości wytrzymałościowe. Zależą one w dużym stopniu od tkaninowego rdzenia taśmy, a w szczególności od rodzaju materiałów, z których jest zbudowany, oraz typu włókien wchodzących w skład jego osnowy (biegnących w kierunku wzdłużnym) i wątku (w kierunku poprzecznym), jako że to właśnie rdzeń przenosi wszystkie naprężenia rozciągające taśmy przenośnika. Tkaninową taśmę transportową uważa się za materiał kompozytowy, składający się z rdzenia (poliamid P i poliamid–poliester EP) oraz warstw wierzchnich (okładek). Koszty taśmy stanowią 10–30% ceny całego przenośnika, dlatego do użycia powinno dopuszczać się jedynie takie taśmy przenośnikowe, które wykazują właściwości określone w odpowiednich normach. Przedmiotem artykułu są zużyte taśmy przenośnikowe z odnowioną górną okładką (odnowione taśmy przenośnikowe). Właściwości wytrzymałościowe wykorzystano do oceny przydatności odnowionych taśm do dalszego wykorzystania w warunkach praktycznych. W artykule przedstawiono analizę właściwości wytrzymałościowych odnowionych taśm tkaninowych w funkcji typu rdzenia. Zaobserwowane wyniki porównano stosując metodę DOE, analizę regresji i korelacji oraz metodę indukcji statystycznej. Wszystkie wnioski uzyskane w oparciu o wyżej wymienione metody są identyczne dla badanych właściwości wytrzymałościowych. Wyniki wskazują, że badane właściwości wytrzymałościowe taśm przenośnikowych nie ulegają istotnej zmianie po renowacji górnej okładki.

Słowa kluczowe: *właściwości wytrzymałościowe, zużyta taśma przenośnikowa, renowacja okładki górnej, typ rdzenia.*

1. Introduction

Belt conveyors have modular designs and are assembled using individual structural components in order to create a transport route of the required length. Such length is limited by the strength of the used type of the belt and by the performance of the drive station. Selecting an appropriate type of belt and possessing the information on its properties may positively affect the conveyor structure and operation.

The external load applied to a conveyor belt induces certain changes in its shape and dimensions; this is referred to as deformation. The changes in shapes or dimensions caused by the effects of an external force are expressed by mechanical properties of conveyor belts.

Mechanical properties are specifically determined in various regulations (technical and technological) and the requirements for conveyor belts are based on particular desired applications. Investigation of these properties is therefore very beneficial for users as well as manufacturers. The investigation thereof has been described by many authors. Kessentini et al. investigated the behaviour of composite polyester/polyamide railway fabric conveyor belts covered by rubber plies after water penetrated inside. They performed hygroscopic gravimetric experiments to assess the moisture diffusion coefficients; gravimetric experiments were performed separately for rubber cover layers, fabric carcasses, and belt specimens [22]. Wong et al. investigated the effects of moisture absorption on the residual tensile strength of uncorrected grooved and double-coated corrected carbon/epoxy composites [29]. Hakami et al. investigated the wear, surface roughness, and temperature increase of styrene-butadiene rubber, natural rubber, and nitrile-butadiene rubber when shifted over abrasive materials of various sizes, with normal load variation. The properties of rubber, such as tensile strength and elongation at break were regarded as input parameters [19]. In another paper, Hakami et al. investigate rubber wear rate and mechanisms and the related affecting parameters based on the data published in the literature. Wear, fatigue and creation of rollers dominate the wear mechanisms that are affected by the load, motion velocity, hardness, and abrasion. Detailed correlations between the affecting parameters and their impact on the rubber wear rate were determined [20]. Alajmi and Shalwan investigated the correlation between mechanical properties of fillers/epoxy composites and their tribological behaviour. Tensile, hardness, wear, and friction tests were conducted for neat epoxy composites, graphite/epoxy composites, and data palm fibre/epoxy composites with or without graphite. The correlation was made between the tensile strength, modulus of elasticity, elongation at break, and hardness, as individual or combined factors, with the specific wear rate and coefficient of friction of composites [1].

Fedorko et al. describe in their papers [6,7,8,9] a novel and still unpublished method of analysing failures of fabric conveyor belts made loaded by the tensile force [15]. This team investigated the properties of smooth conveyor belts by conducting a tensile test aimed at examining the behaviour of internal structures of belt specimens. They applied the metro-tomographic analysis to monitor the behaviour of internal structures of belt specimens exposed to a load. They used the measurements of distances between individual fibres in the warp and weft of the carcass and apply a simple analysis of the extent of the observed defects. The presented results prove that the method of industrial tomography clearly identifies failures of conveyor belt carcasses and facilitates identification of individual fibres, puncture, and separation of individual layers [16,17]. In his paper [18], Fedorko describes the application of FEM-based simulation tools within his research and analyses. Mazurkiewicz claims that extensive, unanalysed reduction of specimen surface during a test significantly affects the accuracy of the a structure model analysed applying the FEM. His article [24] provides the analysis of this problem with regard to modelling the adhesive joining of rubber materials and rubber seal which represents the method to be adopted for identification of strength characteristics of analysed materials, thus eliminating the inaccuracy of the FEM model.

The results are often evaluated by applying the DOE method. This method was used to test the tension of rubber conveyor belts [3] and identify the factors affecting the value of the dynamic impact of a load of conveyor belts [4]. The results of the tests of rubber products, in terms of their quality, aimed at identifying the limit value of impact load, are presented by Ambrisko et al. in the paper [5]. The investigation of the resistance of conveyor belts to punctures and longitudinal slitting is described in the paper by Bajda [9].

Moezzi et al. investigated industrial fabrics used in fabric conveyor belts. They observed an interesting mechanism of degradation of fabrics made of nylon 66 after exposure to the UV radiation at various exposure times; this was confirmed by the results observed for mechanical properties of the specimens. The results of this study may be used in industrial applications of belt conveyors made of nylon 66, exposed to the solar UV radiation [25]. In his paper [11], Barburski analysed mechanical properties of conveyor belts in three main production stages: raw fabric, fabric impregnated with latex, and a conveyor belt. He observed that the differences in mechanical properties of products in individual stages of conveyor belt production depend on the number of the points of intersection in fabric threads.

In his paper [2], Ambrisko describes testing the cover layers of conveyor belts aimed at the identification of abrasion resistance and hardness of cover layers as well as the changes therein caused by natural ageing. The impact of natural ageing on the strength parameters of steel cord conveyor belts are described in the paper by [10] Bajda and Hardygora. In his paper [23] Long et al. describes the investigation of the effects of various factors on the strength of joints in steel-cord conveyor belts; they also examined contact between steel cords and rubber. The paper by Blazej et al. [12] describes the monitoring of creep and stress relaxation occurring in two different types of thermally vulcanised joints in fabric conveyor belts.

Saderova presents in her paper the experimental results of the tensile properties of rubber-fabric composites in relation to the carcass type [27]. Petrikova et al. investigate mechanical properties of conveyor belt specimens in tensile tests conducted at different load speeds; relaxation tests were carried out while measuring also the abrasion index [26]. Harsha investigated the mechanical and tribological properties of various injection-formed polyaryletherketones and their composites. He also examined the tensile strength, modulus of elasticity, elongation at break, flexural strength, flexural modulus, hardness and impact strength [21]. Bojic et al. also tested tensile properties of two plastic woven belts (maximum force, maximum expansion, force at break, and elongation at break) and based on the obtained results they defined a belt that is more suitable for applications in conveyors [14].

The methodology of testing conveyor belts, especially their service life affected by the wear of cover layers, is discussed by authors [28, 8, 13]. The renovated conveyor belts are discussed in papers [6, 7]. These papers present the evaluation of the belt damage degree conducted while applying the logistic regression.

2. Material and methods

2.1. Problem formulation

The fundamental structural components of a conveyor belt include carcass (fabric, steel) and cover layers. Cover layer thickness is affected by the properties of the transported material.

For the purpose of accurate identification of mechanical properties of conveyor belts, a whole conveyor belt (including individual components) must be subjected to tests. Mechanical tests should be conducted and evaluated considering the fact that a conveyor belt is made of several different types of materials. Especially rubber composites, forming the cover and adhesive layers of a conveyor belt, show more complex processes of wearing or deformation during the mechanic stress [20].

The purpose of the experiments was to assess selected mechanical properties of worn fabric conveyor belts with renovated top cover (renovated conveyor belts).

The renovated conveyor belt shall mean a worn belt which has been decommissioned and its top cover layer has been subsequently renovated.

2.2. Conveyor belt specimens

The carcass of a fabric conveyor belt consists of the load-bearing carcass with several fabric plies that are coated with a special rubber compound that ensures adhesion between the plies.

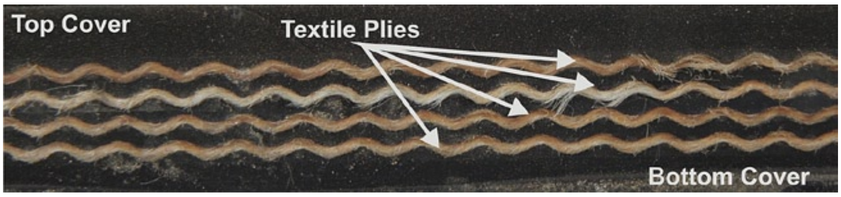


Fig. 1. Fabric conveyor belt (composite)

Carcasses are reinforced with various types of natural and synthetic materials. Most frequently used are conveyor belts with polyamide (P) or polyamide-polyester (EP) carcasses. A polyamide (P) fabric carcass contains polyamide fibres in the weft and in the warp. A polyamide-polyester (EP) carcass consists of polyamide fibres in the weft and polyester fibres in the warp.

- Polyamide is a fully synthetic fibre of good elasticity, high resistance to dynamic stress, resistance to moisture and chemicals. It shows high strength, good fatigue values, but high elongation.
- Polyester is a fully synthetic fibre of good tensile properties, resistance to moisture, acids, and bases. Compared to polyamide, it shows lower strength and lower elongation.

Selected properties of polyamide and polyester are listed in Table 1.

Table 1. Comparison of selected properties of polyamide and polyester

Material	Polyamide (P)	Polyester (E)
Fibre strength [N.mm^2]	740-910	830-970
Breaking force (wet) [%]	65-80	60-70
Specific weight [gcm^{-2}]	1.14	1.38
Elongation at break [%]	12-18	10-15
Elongation at break (wet) [%]	15-25	12-18
Permanent elongation [%]	0.5-3.0	0.3-0.9
Elastic elongation [%]	0.6-1.5	0.5-1.2
Maximum elongation at 1T [%]	1.5-3.0	0.8-2.5

When compared to a polyamide fibre, a polyester fibre shows lower ductility and higher shape stability, and its elongation at use is lower. It is beneficial to use the combination of polyester fibres in the warp and polyamides fibres in the weft as this combination ensures small extension in the longitudinal direction, but good troughability and flexibility of the belt in the transverse direction. When compared to polyamides, polyesters show certain disadvantage- higher density which is manifested by a higher belt weight.

The experiment was carried out using the specimens of fabric conveyor belts with two types of carcasses (P, EP), three different strength values (800; 1,000; and 1,250 N.mm^{-1}), and different numbers of fabric plies (3

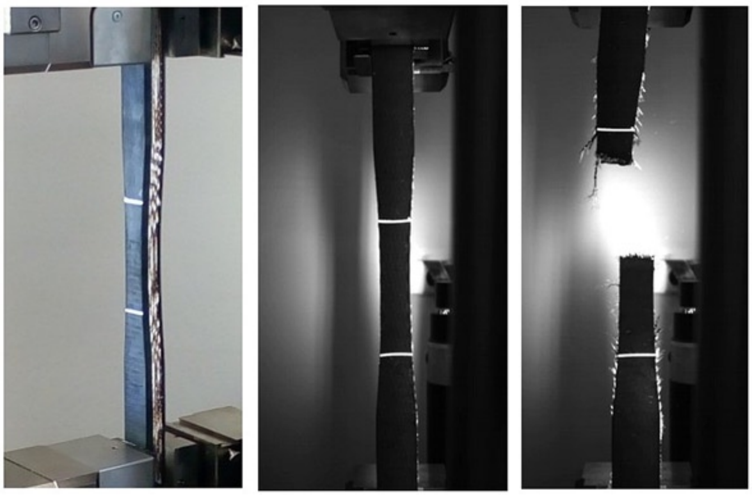


Fig. 3. Fixing and testing the specimen

and 4). 5 specimens were prepared from each type of conveyor belt.

The tested specimens, in the shape of a double-sided blade (Fig. 2), were cut out from the conveyor belt applying the prescribed method. The test pieces were cut out parallelly to the belt axis, in the minimum distance of 50 mm from the belt edge. The specimens have to be cut out perpendicularly to the belt surface and must not contain any joints. The specimen length was 400 mm; the width was 35 mm; and the thickness was determined by the thickness of the conveyor belt from which the particular specimen was cut out. On the prepared specimens, two contrast lines (reference lines −1) were marked in the mutual distance of 100 mm, perpendicularly to the specimen axis. The lines were marked symmetrically to the specimen centre.

Prior to the measurements, the specimens were subjected to conditioning (ISO 18573 standard).

2.3. Experiment execution

The experimental measurements were carried out using the Zwick Roell Z250 testing machine. It is a stationary machine with a linearly growing load. The Z250 machine provides a versatile environment for all kinds of axial tests at lower displacement rates. In the experiment, the testing machine was used in tensile tests. The measurement procedure was as described below.

In the first step, the testing machine was launched together with the operating software for tensile tests for a particular type of the tested materials. In the following step, the test parameters were set, in compliance

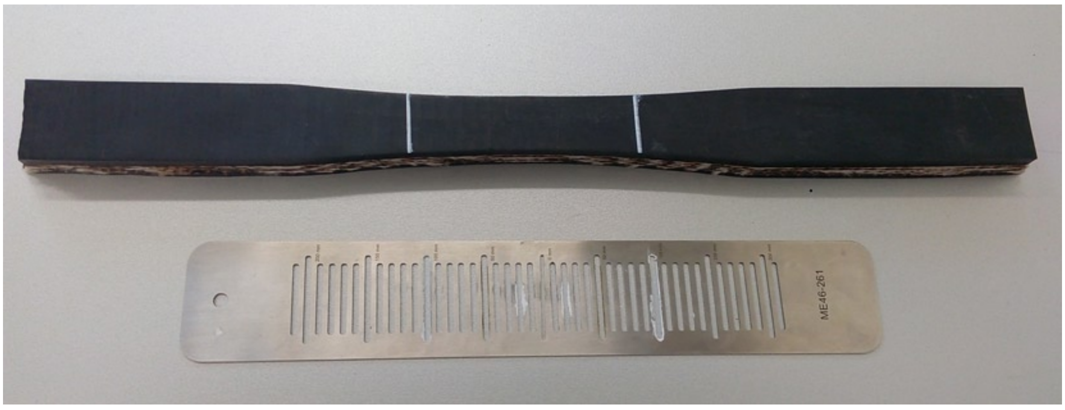


Fig. 2 Test specimen

with the EN ISO 283 standard (velocity of 100 mm.min^{-1}). Subsequently, the specimen was fixed in pneumatic jaws (Fig. 3). The test piece dimensions were entered in the software (measured thickness, width, and tested length). This was followed by performing the measurements that included continuous stretching of the test specimen (until the specimen was exposed to the reference force, or until the specimen rupture) and recording the measured data. After the measurements were completed and the values were recorded, the specimen was removed from the machine and a new specimen was inserted. If a test specimen showed no disturbance between the reference lines, or if the test piece was sliding between the jaws during the test, the specimen was excluded. Each valid measurement with a particular specimen type was repeated three times (Fig. 4).

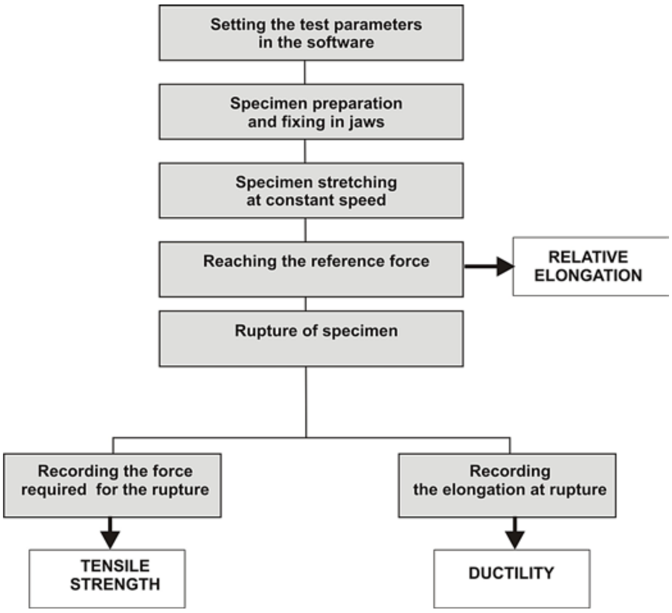


Fig. 4. Measurement scheme

3. Theory/calculation

The strength of conveyor belts necessary for a particular operation is determined using the strength calculation. Total dynamic resistance must be identified for both strands (upper and lower) of the belt conveyor; this equals the total circumferential force on the drive pulley at a stable operation of the conveyor. The magnitude of the total circumferential force can be used to calculate the magnitudes of tensile forces in the belt. They may subsequently be used to identify the necessary belt strength. The calculation of the conveyor belt strength is made while considering its inhomogeneous structure. Considerations are only given to the belt carcass (fibres of the tractive component) because the strengths of other belt parts are relatively small compared to the carcass strength. The strength of the belt is expressed by the force applied to 1 meter of its width because the size of the area of the load-carrying cross-section, especially when a fabric ply is used, can only be identified with high inaccuracy. The belt inspection is carried out while considering only the tensile stress; other stresses are considered by a relatively high safety.

3.1. Methods of experimental research – the test of conveyor belts

The tests of selected mechanical properties (tensile properties) of conveyor belts were carried out pursuant to specified standards. These tests were divided into groups and performed as described in Table 2.

3.1.1. Relative elongation at reference load

The percentage of relative elongation was identified using the following formula:

$$\varepsilon_R = 100\frac{(l_R - l_1)}{l_1} \quad (1)$$

where:

- ε_R is elongation of the test piece [%];
- l_1 is the initial length (initial distance between the reference lines) [mm]; and
- l_R is the length recorded at the reference load of the test piece (final distance between the reference lines) [mm].

3.1.2. Tensile strength

The tensile strength represents the maximum potential stress in the stretched material which the material as a whole is able to resist right before the rupture.

It was calculated using the following formula:

$$F_S = \frac{F_R}{b}, \quad (2)$$

where:

- F_S is the tensile strength [N.mm^{-1}];
- F_R is the force loading the test piece at the moment of rupture [N]; and
- b is the width of the test piece [mm].

3.1.3. Ductility

Ductility is the relative longitudinal elongation of the test piece at the moment of rupture. The percentage of ductility was calculated using the following formula:

$$\varepsilon = 100\frac{(l_2 - l_1)}{l_1}, \quad (3)$$

Table 2. Tensile properties of conveyor belts

Test name	Test mechanism
Relative elongation at reference load	It was performed during the measurement of tensile strength. A test piece was exposed to the tensile force at constant velocity and records were made on the elongation of the working section when reaching the reference load (reference force).
Tensile strength	A test piece was exposed to the tensile force at constant velocity until the rupture occurred and records were made on the required force.
Ductility	It was performed during the measurement of tensile strength. A test piece was exposed to the tensile force at constant velocity until the puncture occurred and records were made on the elongation of the working section at the moment of rupture.

where:

ε is the ductility [%];

l_1 is the initial length (initial distance between the reference lines) [mm];

l_2 is the length recorded right before the rupture of the test piece (final distance between the reference lines) [mm].

3.2. Methods of evaluation and comparison

The Design of Experiments (DOE) method facilitates better understanding and improving technological and laboratory processes. This method enables identification of the process-entering factors with decisive effects on the monitored outputs (output factors or output variables). The assessment of the importance of different effects of factors and identification of optimal conditions were facilitated by the execution of well-prepared experiments. The evaluation of experimental results was carried out using the statistical methods that enable the assessment of whether a change in the monitored output variable was caused by the effects of input factors.

There are different types of experiment planning. The experiment described herein was planned to apply the full three-factorial design in which each considered factor acquired two levels: low „-" and high „+". The effect of the monitored factor on the output variable represented the difference between the average temperatures of the output variable at the low and high levels of the given factor.

The description and comparison of the monitored variables were carried out applying the *basic statistical methods*, including the estimation theory, testing statistical hypothesis, regression and correlation analyses. The predetermined statistical hypotheses were tested by parametrical and non-parametrical tests (one-sample t-test, non-parametric Wilcoxon one-sample test, paired t-test, paired Wilcoxon test, Shapiro-Wilk test for normality, one-way ANOVA test, Kruskal-Wallis test, etc.). The decision-making on accepting or rejecting the null hypothesis was carried out using the p-value. If the p-value was less than the significance level α, then the null hypothesis was rejected. If the p-value was more than, or equal to, the significance level α, the null hypothesis was accepted.

The purpose of the *regression analysis* is to evaluate the existence of relationships between two or more variables and find a suitable regression model for a particular relationship. General considerations were made while presuming a single measurable continuous explanatory (output) variable Y and k explanatory independent variables X_i for $i=1, 2,\dots, k$. The considered conventional multiple regression model was as follows:

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i X_i + \varepsilon \,, \tag{4}$$

where β_0 and β_i for $i=1, 2,\dots, k$ are the regression model parameters and ε is the random error.

The statistical significance of the model (or of individual parameters) was verified applying the F-test of statistical significance of the model (or the test of the statistical significance of the regression parameter). The quality of the regression model was verified using the coefficient of determination R^2. The obtained results were evaluated and compared using the R and Minitab software products.

4. Result and discussion

The experimental research was carried out with the following objectives:

- Compare the tensile properties (relative elongation at the reference load, tensile strength or ductility) of the renovated conveyor belt with the reference values obtained for new conveyor belts.
- Identify which of the input variables (carcass type, number of fabric plies, and conveyor belt strength) affect the output vari-

Table 3. Basic numerical characteristics of output variables

Carcass type	**P**						**EP**					
Strength	**800**		**1,000**		**1,250**		**800**		**1,000**		**1,250**	
Number of plies	**3**	**4**	**3**	**4**	**3**	**4**	**3**	**4**	**3**	**4**	**3**	**4**
Y_1												
Average	1.52	1.40	1.77	1.63	2.09	1.81	1.06	0.90	1.10	0.99	1.27	1.01
Maximum	1.80	1.70	2.20	2.00	2.40	2.20	1.30	1.20	1.40	1.30	1.50	1.30
Minimum	1.30	1.00	1.50	1.30	1.70	1.50	0.80	0.70	0.90	0.80	1.00	0.80
St.deviation	0.14	0.18	0.19	0.20	0.24	0.20	0.14	0.14	0.15	0.14	0.15	0.15
$Yref_1$	1.50	1.41	1.70	1.60	2.10	1.80	1.00	0.92	1.10	1.00	1.20	1.00
Y_2												
Average	906.0	978.2	1195.6	1230.6	1466.9	1517.9	926.6	991.4	1181.3	1236.8	1489.9	1551.9
Maximum	979.0	1029.0	1238.0	1300.0	1533.0	1660.0	996.0	1083.0	1270.0	1334.0	1564.0	1628.0
Minimum	848.0	911.0	1110.0	1169.0	1405.0	1417.0	835.0	880.0	1107.0	1122.0	1433.0	1471.0
St.deviation	33.07	30.88	35.18	37.36	34.88	64.59	41.28	62.30	41.90	62.47	44.21	52.58
$Yref_2$	899.0	988.0	1210.0	1232.0	1455.0	1518.0	942.0	991.0	1190.0	1228.0	1473.0	1559.0
Y_3												
Average	21.71	21.76	22.81	22.86	24.00	24.10	17.81	18.24	18.57	18.62	19.52	19.57
Maximum	23.00	24.00	25.00	25.00	26.00	26.00	20.00	20.00	20.00	20.00	21.00	21.00
Minimum	20.00	20.00	21.00	21.00	22.00	22.00	16.00	16.00	17.00	17.00	18.00	18.00
St.deviation	0.96	1.14	1.40	1.15	0.95	1.14	1.17	1.26	1.08	1.02	0.98	1.08
$Yref_3$	22.00	22.00	22.10	2.50	23.80	24.50	18.20	18.50	18.90	19.00	19.40	19.20

ables (relative elongation at the reference load, tensile strength, or ductility) of the renovated conveyor belt.
- Create regression models of the relationships between the output variables and the input variables of the renovated conveyor belt.
- Compare tensile properties of the renovated conveyor belts with polyamide or polyamide-polyester fabric carcasses.

4. 1. Comparison of tensile properties with the reference values

The experimental research was carried out using 60 test specimens extracted from renovated conveyor belts. The experiment was carried out using renovated rubber-textile conveyor belts with the polyamide or polyester carcass of various strengths (800 $N.mm^{-1}$; 1,000 $N.mm^{-1}$; or 1,250 $N.mm^{-1}$) and with 3 or 4 textile plies. The investigation was concentrated on three basic tensile properties (output variables): relative elongation at the reference load Y_1 [%], tensile strength Y_2 [Nmm^{-1}], ductility Y_3 [%].

The reference value $Yref_z$, z=1,2,3 shall mean the value obtained from the measurements of the investigated properties Y_z (z=1,2,3) for three specimens of new (unused) conveyor belts. The comparison of tensile properties of specimens of renovated conveyor belts with the reference values was carried out applying the basic methods of statistical induction - testing statistical hypotheses. Basic numerical characteristics of the investigated properties of renovated conveyor belts (the output variables) and the relevant reference values are listed in Table 3.

The normality of the data set of renovated conveyor belts was verified applying the Shapiro-Wilk test of normality. The null hypothesis is that the sample come from a Normal distribution. If p-value < α, the null hypothesis is rejected. For output variable Y_1 (relative elongation) and Y_2 (tensile strength), the normality requirement was met in all the cases (Table 4).

The comparison of the measured tensile properties with reference values was carried out applying the parametric one-sample t-test or non-parametric Wilcoxon one-sample test (in the case that the normality requirement is not met).

The results of the comparison indicate that there are no statistically significant differences between the measured values of tensile properties (Y_z , z=1,2,3) and the reference values $Yref_z$, z=1,2,3 (p-value>α). The resulting p-values of the tests are listed in Table 5.

The following sections will deal with the values obtained exclusively from the specimens of renovated rubber-textile conveyor belts.

4.2. DOE method

The purpose of the experiment was to identify which is the input factors (conveyor belt strength A, number of fabric plies B and carcass type C) have statistically significant effects on the value of the output factor Y_z, z=1,2,3 (relative elongation at the reference load Y_1 [%], tensile strength Y_2 [Nmm^{-1}], ductility Y_3 [%]).

All input factors were tested at two different levels. The experiment was carried out using two types of renovated rubber conveyor belts (polyamide and polyester) with the strengths of 800 $N.mm^{-1}$ or 1,250 $N.mm^{-1}$ and with 3 or 4 fabric plies. The levels of individual factors are listed in Table 6.

Each experiment was conducted twice and the evaluation was carried out using the average value of the output variable. The experiment was evaluated considering three main factors (A, B and C) and second-order interactions (AB, AC, and BC). The effects of the main factors and interactions for all three output variables are listed in Table 7. The significance of the individual effects of the factors or interactions on the output variables was tested by the t-test and by the determination of the p-value. Statistically significant effects on the output variable were observed for those main factors, or interactions, for which the p-value < 0.05*.

Table 4. Shapiro-Wilk test of normality (α=0.05)

Carcass type	P						EP					
Strength	800		1,000		1,250		800		1,000		1,250	
Number of plies	3	4	3	4	3	4	3	4	3	4	3	4
Y_1												
p-value	0.259*	0.064*	0.219*	0.193*	0.105*	0.159*	0.188*	0.076*	0.060*	0.113*	0.074*	0.115*
Y_2												
p-value	0.600*	0.212*	0.142*	0.698*	0.145*	0.243*	0.687*	0.190*	0.906*	0.090*	0.060*	0.061*
Y_3												
p-value	0.006	0.092*	0.010	0.081*	0.071*	0.012	0.037	0.082*	0.002	0.015	0.019	0.004

*Note: *p-value> α*

Table 5. Test results (one sample t-test, Wilcox one-sample test, α=0.05)

Type of carcass	P						EP					
Strength	800		1,000		1,250		800		1,000		1,250	
Number of plies	3	4	3	4	3	4	3	4	3	4	3	4
Y_1												
p-value	0.530*	0,906*	0.105*	0.510*	0.784*	0.831*	0.070*	0.534*	0.882*	0.651*	0.055*	0.673*
Y_2												
p-value	0.345*	0,161*	0.075*	0.863*	0.135*	0.997*	0.102*	0.978*	0.352*	0.528*	0.096*	0.541*
Y_3												
p-value	0.177*	0,348*	0.080*	0.171*	0.345*	0.203*	0.132*	0.413*	0.673*	0.103*	0.569*	0.232*

*Note: *p-value> α*

Table 6. Levels of main input factors

Factor level	Main input factors		
	Strength [N.mm^{-1}]	Number of fabric plies [m]	Carcass type
	A	B	C
Low (-)	800	3	P
High (+)	1,250	4	EP

The analysis of the results indicates that all three main factors (A, B, and C) have statistically significant effects on the relative elongation (Y_1) (p-value < α).

Two main factors: factor A (strength) and factor B (number of fabric plies) have statistically significant effect on the tensile strength (Y_2).

Two main factors: factor C (carcass type) and factor A (strength) have statistically significant effects on ductility (Y_3).

4.3. Regression and correlation analysis

Monitoring the effects of the input variables (conveyor belt strength, number of fabric plies and carcass type) on the output variables was based on the following model:

$$Y_z = f\left(STRENGTH, NUMBER\,OF\,PLIES,\,CARCASS\,TYPE\right), \text{ for } z=1,2,3 \quad (5)$$

where Y_1 is the relative elongation at the reference load [%], Y_2 is the tensile strength [Nmm^{-1}], and Y_3 is the ductility [%].

The point estimate of the regression model has the following form:

$$Y_z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon \text{ , for } z=1,2,3 \text{ ,} \quad (6)$$

where $\beta_0, \beta_i, i=1,2,3$ are the parameters of the models; X_1 is the strength; X_2 is the number of fabric plies; and X_3 is the carcass type (X_3=0, if the carcass type is P, X_3=1, if the carcass type is EP). Mutual relationships were identified considering the average values of individual variables. The model parameters were identified applying the method of least squares.

The most optimal regression model of the relationship between the variable Y_1 and the input variables was the following model (type I):

$$Y_1 = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3. \quad (7)$$

The analysis of the results indicated that all input variables (X_1, X_2 and X_3) have statistically significant effects on the relative elongation (Y_1) (p-value < α).

For the variable Y_2, the best regression model is as follows (type II):

$$Y_2 = b_0 + b_1 X_1 + b_2 X_2. \quad (8)$$

Two input variables have statistically significant effects on the tensile strength (Y_2): the strength (X_1) and the number of fabric plies (X_2).

As to the variable Y_3, the best regression model is as follows (type III):

$$Y_3 = b_0 + b_1 X_1 + b_3 X_3. \quad (9)$$

The analysis of the results indicated that the strength (X_1) and the carcass type (X_3) have statistically significant effects on ductility (Y_3).

The point estimates of regression model parameters, statistical significance of the parameters (p-value), coefficient of determination, and p-value of the model are listed in Table 8.

All parameters of the examined regression models, as well as the given regression model, are statistically significant (p-value < α). The values of coefficients of determination R^2 indicate that the quality of the regression model is high.

Table 7. Effects of the main factors and interactions of second-order (α =0.05)

Output variable		A	B	C	AB	AC	BC
Y_1 - Relative elongation	effect	0.320	-0.205	-0.640	-0.065	-0.160	-0.005
	p-value	0.030*	0.046*	0.015*	0.144	0.060	0.795
Y_2 - Tensile strength	effect	556.10	62.52	22.66	-5.98	5.78	0.88
	p-value	0.005*	0.047*	0.127	0.416	0.426	0.878
Y_3 - Ductility	effect	1.905	0.145	-4.120	-0.070	-0.385	0.095
	p-value	0.040*	0.440	0.019*	0.664	0.192	0.571

Table 8. Point estimates of regression models

Output variable	Model parameters				R^2	p-value
	b_0	b_1	b_2	b_3		
Y_1 - Relative elongation	1.545	0.001	-0.162	-0.645	95.44	2.10^{-10*}
p-value	2.10^{-4*}	1.10^{-3*}	2.10^{-2*}	2.10^{-6*}		
Y_2 - Tensile strength	-216.622	1.234	52.758	-	99.48	2.10^{-11*}
p-value	9.10^{-4*}	6.10^{-12*}	5.10^{-4*}	-		
Y_3 - Ductility	18.538	0.004	-	-4.15	99.30	2.10^{-10*}
p-value	2.10^{-12*}	6.10^{-7*}	-	1.10^{-10*}		

*Note: * p-value < α*

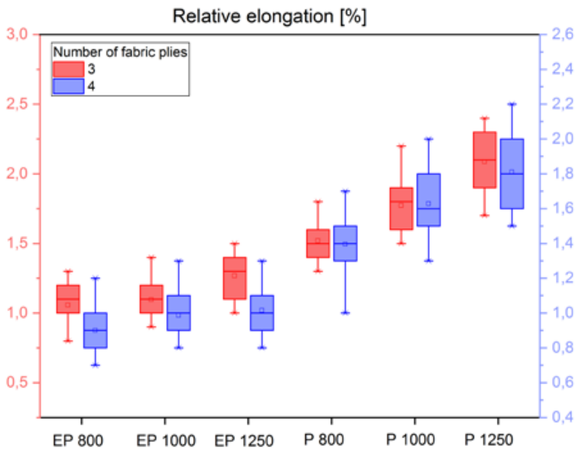


Fig. 5 Boxplot – Relative elongation

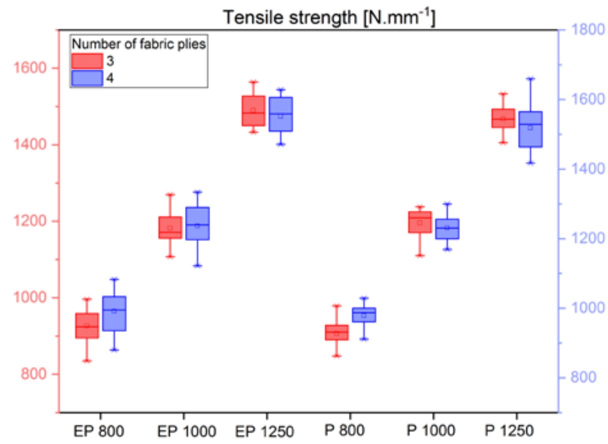


Fig. 7 Boxplot – Ductility

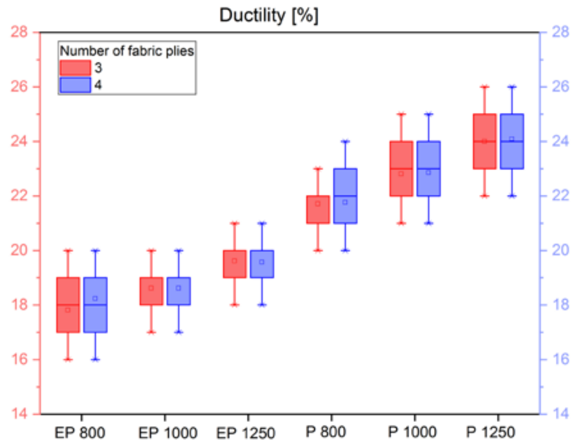


Fig. 6 Boxplot – Tensile strength

4.4. Comparison of tensile properties

This section presents the comparison of tensile properties obtained from the specimens of renovated rubber-textile conveyor belts with the polyamide carcass and the specimens with the polyamide-polyester textile carcass. The basic numerical characteristics of the investigated properties are listed in Table 3. Graphical representations of data sets are shown in Fig. 5 – 7.

The paired comparison of the measured values of individual output variables was carried out applying the paired t-test, or the non-parametric Wilcoxon paired test. The comparison of the three data sets was carried out applying the ANOVA method, or the Kruskal-Wallis test. The test results are listed in Tables 9 to 11. The statistic testing was focused on the following three areas:

- I - effects of the number of fabric plies on the values of individual output variables, on condition that carcass types and strengths are identical (paired t-test or Wilcoxon paired test);
- II - effects of the carcass type on the values of individual output variables, on condition that the strengths of conveyor belts and numbers of fabric plies are identical (paired t-test or Wilcoxon paired test); and
- III - effects of the strength on the values of individual output variables, on condition that carcass types and numbers of fabric plies are identical (ANOVA or Kruskal-Wallis test).

The analysis of the results of Testing I (Table 9) indicates that with identical carcass types (either P or EP) and identical conveyor belt strengths, the number of fabric plies (3 or 4) has a statistically significant effect on the output variable Y_1 (relative elongation) and the output variable Y_2 (tensile strength). On the other hand, it seems that the number of fabric plies has no statistically significant effect, under the given conditions, on the output variable Y_3 (ductility).

The analysis of the results of Testing II (Table 10) indicates that with identical conveyor belt strengths and identical numbers of fabric plies, a carcass type (either P or EP), has a statistically significant effect on the output variable Y_1 (relative elongation) and the output

Table 9. Results of Testing I – number of fabric plies (α =0.05)

Variable	p-value	P800/3 P800/4	P1000/3 P1000/4	P1250/3 P1250/4	EP800/3 EP800/4	EP1000/3 EP1000/4	EP1250/3 EP1250/4	Stat. significance
Y_1	p-value	2.10^{-2}	2.10^{-2}	2.10^{-4}	8.10^{-4}	2.10^{-2}	2.10^{-4}	**Stat. significant**
Y_2	p-value	7.10^{-9}	3.10^{-3}	3.10^{-3}	3.10^{-4}	2.10^{-3}	2.10^{-4}	**Stat. significant**
Y_3	p-value	0.969	0.907	0.646	0.293	0.927	0.970	Stat. insignificant

Table 10. Results of Testing II – carcass type (α =0.05)

Variable	p-value	P800/3 EP800/3	P800/4 EP800/7	P1000/3 EP1000/3	P1000/4 EP1000/4	P1250/3 EP1250/3	P1250/4 EP1250/4	Stat. significance
Y_1	p-value	1.10^{-13}	7.10^{-12}	4.10^{-15}	2.10^{-14}	4.10^{-15}	2.10^{-16}	**Stat. significant**
Y_2	p-value	0.083	0.392	0.239	0.699	0.069	0.070	Stat. insignificant
Y_3	p-value	2.10^{-8}	5.10^{-8}	2.10^{-8}	3.10^{-8}	2.10^{-8}	2.10^{-8}	**Stat. significant**

Table 11. Results of Testing III – strength (α=0.05)

Variable	p-value	P800/3 P1000/3 P1250/3	P800/4 P1000/4 P1250/4	EP800/3 EP1000/3 EP1250/3	EP800/4 EP1000/4 EP1250/4	Stat. significance
Y_1	p-value	1.10^{-3}	2.10^{-8}	4.10^{-5}	3.10^{-2}	**Stat. significant**
Y_2	p-value	4.10^{-51}	2.10^{-42}	7.10^{-46}	2.10^{-37}	**Stat. significant**
Y_3	p-value	1.10^{-6}	2.10^{-6}	3.10^{-5}	2.10^{-3}	**Stat. significant**

variable Y_3 (ductility). On the other hand, it seems that the carcass type has no statistically significant effect, under the given conditions, on the output variable Y_2 (tensile strength).

The ANOVA method (for variables Y_1 and Y_2) and the Kruskal-Wallis test (for variable Y_3) confirmed that renovated conveyor belt strength has a statistically significant effect on all output variables (Table 11).

5. Conclusion

Experimental measurements conducted in laboratory conditions provided the information on tensile properties (relative elongation, tensile strength, and ductility) of selected types of worn conveyor belts with the renovated top cover. Based on the results obtained applying the methods of statistical induction it is possible to assume that the measured values are comparable to the reference values obtained with new, unused conveyor belts with the same carcass type.

Within the research, the output variables (relative elongation, tensile strength, and ductility) were used to compare tensile properties of renovated polyamide and polyamide-polyester fabric conveyor belts. The investigated renovated conveyor belts were of various strengths, as determined by their manufacturers, and various numbers of fabric plies. The comparison was carried out using three input variables (carcass type, belt strength, and the number of fabric plies in a belt) and their effects on three output variables (relative elongation, tensile strength, and ductility).

Tensile properties of renovated conveyor belts were compared applying three different methods and all of them brought the same results.

The conclusions made applying the DOE method, regression analysis, and statistical induction method were as follows:

- The resulting values of relative elongation (Y_1) are affected by the strength of a composite, the number of fabric plies in a composite, and the carcass type.
- The resulting values of tensile strength (Y_2) are affected by the strength of a composite and the number of fabric plies in a composite.
- The resulting values of ductility (Y_3) are affected by the strength of a composite and the carcass type.

Based on all the observed results it may be stated that the carcass type (polyamide or polyamide-polyester) has a statistically significant effect on the resulting values of tensile properties of renovated conveyor belts only for relative elongation and ductility. The type of a composite has no statistically significant effect on the resulting values of the third examined mechanical property of conveyor belts, i.e., the tensile strength.

The analysis of the research described herein indicates that the renovation of the top cover layer of conveyor belts preserves, to a considerable extent, tensile properties of conveyor belts. This information may be helpful for belt users when making decisions on the replacement or renovation of worn conveyor belts.

Acknowledgements

This article was prepared with the support from the project titled KEGA 002TUKE-4/2019 „Modern technologies and innovations in the education of physics at FMT TUKE“ and project titled APVV-18-0248 „Smart belt conveyors “

References

1. Alajmi M, Shalwan A. Correlation between mechanical properties with specific wear rate and the coefficient of friction of graphite/epoxy composites. Materials 2015; 8: 4162-4175, https://doi.org/10.3390/ma8074162.
2. Ambrisko L. Determination of the abrasion resistance and the hardness of rubber covering layers. Albena, Bulgaria, 2 July - 8 July 2017, paper no. 142873: 255-262. SGEM.
3. Ambrisko L, Marasova D. and Grendel P. Determination the effect of factors affecting the tensile strength of fabric conveyor belts. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18: 110-116, https://doi.org/10.17531/ein.2016.1.14.
4. Ambrisko L, Marasova D. Evaluation the quality of rubber composites using the DOE method. Quality - Access to Success 2017; 18: 60-63.
5. Ambrisko L, Taraba V, Szabo S. and Leco M. Belt damage aspect to impact loading. Applied Mechanics and Materials 2014; 683: 102-107, https://doi.org/10.4028/www.scientific.net/AMM.683.102.
6. Andrejiova M, Grincova A, Marasova D. Failure analysis of rubber composites under dynamic impact loading by logistic regression. Engineering Failure Analysis 2018; 84: 311-319, https://doi.org/10.1016/j.engfailanal.2017.11.019.
7. Andrejiova M, Grincova A, Marasova D, Fedorko G, Molnar V. Using logistic regression in tracing the significance of rubber-textile conveyor belt damage. Wear 2014; 318: 145-152, https://doi.org/10.1016/j.wear.2014.06.026.
8. Bahke E A. New developments in belt conveying technology. Bulk Solids Handling 1986; 6: 19-25.
9. Bajda M. Laboratory tests of conveyor belt parameters affecting its lifetime. Albena, Bulgaria, 29 June - 5 July 2017, paper no. 130787: 495-502. SGEM.
10. Bajda M, Hardygora M. The influence of natural ageing processes on the strength parameters of steel cord conveyor belts. International Journal of Mining, Reclamation and Environment 2018; 32: 430-439, https://doi.org/10.1080/17480930.2018.1436030.
11. Barburski M. Analysis of the mechanical properties of conveyor belts on the three main stages of production. Journal of Industrial Textiles 2014; 45: 1322-1334, https://doi.org/10.1177/1528083714559567.
12. Blazej R, Bajda M, Hardygora M. Monitoring creep and stress relaxation in splices on multiply textile rubber conveyor belts. Acta Montanistica Slovaca 2017; 22: 116-125.

13. Blicblau A S. Evaluation of a rubber abrasion test method. Bulk Solids Handling 1984; 4: 353-362.
14. Bojic N, Nikolic R, Banic M, Hadzima B. Evaluation of mechanical properties ofthe two PVC conveyor belts. Communications - Scientific Letters of the University of Zilina 2018; 20: 47-51.
15. Fedorko G, Molnar V, Michalik P, Dovica M, Toth T, Kelemenova T. Extension of inner structures of textile rubber conveyor belt - Failure analysis. Engineering Failure Analysis 2016; 70: 22-30, https://doi.org/10.1016/j.engfailanal.2016.07.006.
16. Fedorko G, Molnar V, Michalik P, Dovica M, Kelemenova T, Toth T. Failure analysis of conveyor belt samples under tensile load. Journal of Industrial Textiles 2019; 48: 1364-1383, https://doi.org/10.1177/1528083718763776.
17. Fedorko G, Molnar V, Dovica M, Toth T, Fabianova J, Strohmandl J, Neradilova H, Hegedus M, Belusko M. Analysis of defects in carcass of rubber-textile conveyor belts using metrotomography. Journal of Industrial Textiles 2018; 47: 1812-1829, https://doi.org/10.1177/1528083717710712.
18. Fedorko G, Molnar V, Honus S, Belusko M, Tomaskova M. Influence of selected characteristics on failures of the conveyor belt cover layer material. Engineering Failure Analysis 2018; 94: 145-156, https://doi.org/10.1016/j.engfailanal.2018.07.034.
19. Hakami F, Pramanik A, Islam N, Ridgway N. Study of effective parameters on wear behavior of rubbers based on statistical methods. Polymers for Advanced Technologies 2019; 30: 1415-1426, https://doi.org/10.1002/pat.4574.
20. Hakami F, Pramanik A, Ridgway N, Basak AK. Developments of rubber material wear in conveyer belt system. Tribology International 2017; 111: 148-158, https://doi.org/10.1016/j.triboint.2017.03.010.
21. Harsha AP, Tewari US. Tribo performance of polyaryletherketone composites. Polymer Testing 2002; 21: 697-709, https://doi.org/10.1016/S0142-9418(01)00145-3.
22. Kessentini R, Klinkova O, Tawfiq I, Haddar M. Modeling the moisture diffusion and hygroscopic swelling of a textile reinforced conveyor belt. Polymer Testing 2019; 75: 159-166 https://doi.org/10.1016/j.polymertesting.2019.01.013.
23. Long X, Li X, Long H. Analysis of influence of multiple steel cords on splice strength. Journal of Adhesion Science and Technology 2018; 32: 2753-2763, https://doi.org/10.1080/01694243.2018.1505806.
24. Mazurkiewicz D. Problems of identification of strength properties of rubber materials for purposes of numerical analysis. Archives of Civil and Mechanical Engineering 2010; 10: 69-84, https://doi.org/10.1016/S1644-9665(12)60131-9.
25. Moezzi M, Yekrang J, Ghane M, Hatami M. The effects of UV degradation on the physical, thermal, and morphological properties of industrial nylon 66 conveyor belt fabrics. Journal of Industrial Textiles 2019 (article in press), https://doi.org/10.1177/1528083718825316.
26. Petrikova L, Marvalova B, Tuan HS, Bocko P. Experimental evaluation of me2hanical properties of belt conveyor with textile reinforcement and numerical simulation of its behaviour. San Sebastian, Spain, 25 June - 28 June 2013, paper no. 97264: 641-644, https://doi.org/10.1201/b14964-116.
27. Saderova J. Laboratory research of a conveyor belts with a textile carcass for mine conveying. Albena, Bulgaria, 29 June - 5 July 2017, paper no. 142873, pp. 537-544. SGEM.
28. Wolpers F M, Hager M. Tests on the wear behaviour of conveyor belts. Bulk Solids Handling 1990; 10.
29. Wong K, Gong XJ, Aivazzadeh S, Tamin MN. Moisture effects on patch bonded composite repairs. Key Engineering Materials 2016; 709: 3-6, https://doi.org/10.4028/www.scientific.net/KEM.709.3.

Miriam ANDREJIOVA
Department of Applied Mathematics and Informatics
Faculty of Mechanical Engineering
Technical University of Kosice
Letna 9, 042 00 Kosice, Slovak Republic

Anna GRINCOVA
Department of Mathematics and Theoretical Informatics
Faculty of Electrical Engineering and Informatics
Technical University of Kosice
Letna 9, 042 00 Kosice, Slovak Republic

Daniela MARASOVA
Logistics Institute of Industry and Transport
Faculty of Mining, Ecology, Process Control and Geotechnology
Technical University of Kosice
Letna 9, 042 00 Kosice, Slovak Republic

E-mails: miriam.andrejiova@tuke.sk, anna.grincova@tuke.sk, daniela.marasova@tuke.sk

Article citation info:
STETTER R, GÖSER R, GRESSER S, TILL M, WITCZAK M. Fault-tolerant design for increasing the reliability of an autonomous driving gear shifting system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 482–492, http://dx.doi.org/10.17531/ein.2020.3.11.

Ralf STETTER
Richy GÖSER
Sebastian GRESSER
Markus TILL
Marcin WITCZAK

FAULT-TOLERANT DESIGN FOR INCREASING THE RELIABILITY OF AN AUTONOMOUS DRIVING GEAR SHIFTING SYSTEM

PROJEKTOWANIE TOLERUJĄCE USZKODZENIA ZWIĘKSZAJĄCE NIEZAWODNOŚCI SYSTEMU ZMIANY BIEGÓW POJAZDU AUTONOMICZNEGO

The reliability of technical systems can be greatly reduced if possible faults cannot be accommodated but lead to system shut-down with sometimes catastrophic consequences. The algorithms and systems of fault-tolerant control were developed in the last years into a powerful tool to accommodate such faults. Additionally, it became obvious that the design of a technical system can ease or hinder the application of these tools and can also lead to the accommodation of faults be itself. This kind of design – fault-tolerant design – and its components are presented in this paper on the example of a shifting system for the gear box an autonomous driving race car. This race car competes in the well-known formula student driverless competition; in such competitions the reliability of the car and the capability to accommodate not avoidable faults is of paramount importance. The different elements of fault-tolerance incorporated in the design of the gear shifting system are explained on the basis of an established model of product concretization.

Keywords: *fault-tolerant design, design methods, design for reliability, automated gear shifting.*

Niezawodność systemów technicznych może być znacznie ograniczona w przypadku braku odpowiedniej akomodacji uszkodzeń, która może doprowadzić do awarii systemu mogącej mieć katastrofalne konsekwencje. W celu przeciwdziałania temu niepożądanemu zjawisku, w ostatnich latach opracowano szereg algorytmów sterowania tolerującego uszkodzenia, umożliwiających odpowiednią akomodację uszkodzeń. Dodatkowo, oczywistym jest, że sposób projektowania danego systemu może ułatwić lub utrudnić funkcjonowanie powyższych algorytmów. Może one również sam w sobie umożliwiać odpowiednią akomodację uszkodzeń. Tak sposób projektowania, projektowanie tolerujące uszkodzenia, jest przedmiotem niniejszej pracy na przykładzie systemu zmiany biegów w autonomicznych pojeździe wyścigowym. Powyższy pojazd współzawodniczy w znanych studenckich zawodach wyścigowych pojazdów autonomicznych. Oczywistym jest fakt, że w tego typu zawodach, niezawodność pojazdu i jego zdolność akomodacji uszkodzeń jest szczególnie ważna. W pracy rozważa się różne element projektowania tolerującego uszkodzenia systemu zmiany biegów opisanego na podstawie ustalonego modelu konkretyzacji produktu.

Słowa kluczowe: *projektowania tolerujące uszkodzenia, metody projektowania, projektowanie dla niezawodności, automatyczna zmiana biegów.*

1. Introduction

This paper is focusing on the methods and tools of fault-tolerant design applied to a gear shifting system. This application is explained and reflected for a concrete representation of this concept. The main objective of the strategies, methods, tools and algorithms as well as general insights of fault-tolerant design is the support of engineers in the development of technical systems, which are fault-tolerant as a consequence of their controllability and diagnosability but also as a consequence of their inherent fault-tolerant design qualities [50]. Reliability is the probability that a technical system is able to perform its intended function for a specified period of time under specified operating conditions (e.g. loads, temperatures) [43]. Current research aimed at increasing the reliability of technical systems is addressing dynamic analyses of multi-state systems [27], reliability management systems [35], non-direct determination of system deterioration [54] fuzzy logic for vulnerability assessment [55] and structure learning algorithms [29]. The research outcomes with this objective were already successfully applied in many cases such as in a reliability estimation for momentum wheel bearings [24], for load-sharing failures in parallel systems [61], the level adjustment of quadruped robots [13] and the operational reliability of rail vehicles [33]. The development of reliable cognitive technical systems presents a continuous challenge for research teams [40]. A new direction is the evaluation of the reliability of technical systems already in the conceptual design phase based on effect chains [8].

It is important to note that it is impossible for current complex technical systems to avoid faults completely. The accommodation of these faults can be a decisive approach to increase the reliability of the systems. Firstly, this accommodation can be based on the numerous research works concerning fault-tolerant control (a concise overview is given by Blanke et al. [6]). Current developments in this area include adaptive control schemes that are robust to parameter uncertainties, disturbances and saturation [1], virtual sensors based on quadratic boundedness [51] and adaptive sliding mode observer based fault-tolerant control [37]. Zhang et al. also propose tolerance measures for systems with sensor failures [60]. An important prerequisite for active fault-tolerant control are elaborate fault detection and fault diagnosis algorithms (compare [11]). In recent years, several novel fault detection algorithms were proposed such as time synchronous resample and adaptive variational mode decomposition [62] or convolutional neural networks with global average pooling [30].

Secondly, the emerging field of fault-tolerant design can be employed. It is important to note that this field is strongly connected with fault-tolerant control, as fault-tolerant design can be an important enabler for these algorithms and systems. These two concepts in combination have the potential to enhance to fault-tolerance and with this the reliability on many levels.

As mentioned above, faults cannot be totally avoided in complex technical systems. They can be defined as an unwanted deviation of at least one parameter or property of the respective technical system from the satisfactory, regular condition (compare [28]), e.g. a sensor malfunction leading to the inability of this sensor to measure an important system parameter. Faults can be distinguished from failures, which can be defined as permanent interruptions of the capability of a product to perform the planned functionality; a failure indicates a complete breakdown of the product. Fault-tolerant technical systems are able to accommodate certain faults; i.e. to enable a degraded but still satisfactory performance of the system in the case of a fault. The main contribution of this paper is an in-depth exploration of design aspects which can increase the fault-tolerance of technical systems. On the one hand, these design aspects enable and ease fault-tolerant control. On the other hand, these design aspects increase the fault-tolerance independently, e.g. by adding redundant elements. These activities can be summaries with the notion "fault-tolerant design". The paper provides a novel structure for fault-tolerant design based on levels of product concretization, presents several concrete measures – design aspects – for enhancing the fault-tolerance and explains these measures on a real-life case – the design of a gear-shifting system of a race-car. Concrete design characteristics are described in this paper, which either increase the controllability or diagnosability of the gear shifting system or directly increase its fault-tolerance. In order to have a background for the detailed explanation of these design characteristics, the next section describes the essence of fault-tolerant design and presents a model to structure the later discussion. The newly developed gear shifting system for the formula student driverless competition is explained in Section 3. Section 4 discusses the distinct design characteristics. The last section concludes the discussion and gives an outlook on future research activities.

2. Fault-tolerant design

Faults are unwanted deviations of one or more parameters or properties of a technical system, which may lead to a considerable reduction of the reliability of this system. It is possible to differentiate faults analysing their behaviour over time. Permanent faults are sudden alterations, which lead to an on-going change of physical parameters of the technical system or its structures [56]. Drift-like faults are evolving over time, but lead after a certain delay also to an on-going change of physical parameters of the technical system or its structures. Intermittent faults are deviations which appear and disappear again and again and exhibit commonly a rather short duration. The different behaviour over time is shown in Figure 1.

The quality of a system to allow the accommodation of consequences of any kind of fault and to ensure a still safe operation with only slightly degraded performance may be subsumed under the notion "fault-tolerance" (compare Rouissi and Hoblos [47] and Dubrova [15]). In recent years, especially active fault-tolerant control approaches have exhibited convincing performance, which use a

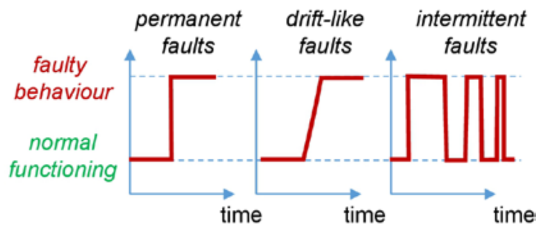


Fig. 1. Faults - behaviour over time (compare [50])

fault detection and identification system and realise the fault handling based on the results of this system [59].

Fault-tolerant design can, on the one side, ease active fault-tolerant control approaches, because it aims, amongst others, at enabling far-reaching monitoring possibilities, thus supporting fault detection and identification activities. On the other side, fault-tolerant design can strengthen the inherent fault-tolerant design characteristics of a technical system, for instance through the application of robust physical effects (compare [50]). Rouissi and Hoblos [47] underline the significance of fault-tolerant design; they highlight that the capability of a system to accommodate faults is directly linked to the design quality. As stated above, this capability has an enormous influence on the reliability of a complex technical system. So far, only a small amount of research is directly concerning fault-tolerant design. A future methodology may be built on the vast amount of research concerning systematic design and product development (Ehrlenspiel and Meerkamm [15], Ponn and Lindemann [44], Pahl et al. [42]). The research outcomes concerning "Design for Monitoring" (DfM), "Design for Control" (DfC) and "Design for Diagnosis" (DfD) may also serve as a basis for building this methodology. The research results concerning "computational design synthesis" (CDS) [38], the application of graph based design languages based on UML [58], the development and modelling of cyber-physical systems [63], a development methodology for mechatronic systems based on SysML [3] and safety analysis based on SysML [36] may expand this basis and allow a digital integration in current engineering processes.

Up to today, a rather small number of research activities can be identified, which are already directly focusing on fault-tolerant design. Additionally, they usually only cover a specific field. Oh et al. [41]) are proposing, amongst others, voting logic and redundant actuation devices in order to increase the fault-tolerance of nuclear power plants. Further authors aim to increase the fault-tolerance of technical systems: of wide-area networks [31], of wireless sensor networks [49], of microelectronics [26] and of frequency converters [57]. The applied approaches include analysis methods such as "reliability modelling" and "redundancy analysis" [57] and design proposals such as "Triple Modular Redundancy" (an intended design of a part of a chip is copied twice and a voter always choses the outcome of at least two designs) [26]. A novel fault-tolerant strategy for distributed actuators is the application of a consensus protocol for fuzzy multi-agent systems; this is described by Chen et al. [9]. In automotive applications, systems such as active steering systems are modelled and analysed in order to evaluate their fault-tolerance [20]. Similar modelling approaches were also applied to rail systems [14]. Fault-tolerant design is also addressed for improving the reliability of airplanes. A fault-tolerant interior permanent motor for an electrical power steering is presented by Bianchi et al. [5]. Brando et al. [7] propose a fault-tolerant design of redundant axial-flux motors for the electric steering of aircraft nose landing gears. The research of Nie et al. [39] focuses on the fault-tolerant design of electronic power converters, this work is continued by Tian et al. [53]. Additional challenges are distributed control structures for electric transformers; fault-tolerant control structures are addressed by Saeed et al. [48]. An adjacent field concerns damage-tolerant design [52]; the point of main emphasis is on fatigue and fracture investigations in computational mechanics, but some damage-tolerance approaches have a potential to be integrated in a fault-tolerant design methodology. Another adjacent field is system reliability design and some approaches such as the active disturbance rejection control (ADRC) method (early approaches reported by Gao [21] and Han [22], current applications e.g. by Ramirez-Neria et al. [45]) also dispose of considerable potential to increase the fault-tolerance of technical systems.

It can be concluded that several elements for a fault-tolerant design methodology are already existing. However, only the implementation of elements and tools is not sufficient for the development of performant, reliable and safe mechatronic systems [3]. It is important to note that the early stages of product development are essential and that currently there are no common methods with computational support for these stages [8]. Additionally, an increase of the tolerance of technical systems requires a deep understanding of potential failure modes [52]. Further challenges arise with the distributed control schemes of modular systems; the design of distributed control structures is not trivial [48] and a conscious design of the modular structure is inevitable. It is consequently of vital importance for the development of knowledge and the support of design and control engineers to investigate the models, algorithms and methods which allow a conscious and holistic fault-tolerant design. These models, algorithms and methods need to connect the different phases of design, the different levels of abstraction and the disciplines [14]; through this they can then contribute to increasing the reliability of technical systems.

An initial defining structure for this future methodology was so far presented by Stetter [50]. It is based on the well-known models of product concretization which are widely used in systematic design and product development (compare e.g. Ponn and Lindemann [44]). The four levels of this model are shown in Figure 2.

The basis for any successful development is the requirement level. Requirements are the objectives, goals and specifications which describe the functionality and intended or required characteristics of the technical system under development. In industrial system development processes, requirements are a decisive factor (compare e. g. Bernard and Irlinger [1]) and nearly fifty percent of the top risks are connected to this factor (Hruschka [25]). Like other characteristics and functionalities of a technical system, also the intended fault-tolerance should be defined in the early stages of a project. This definition should include the intended level of fault-tolerance as well as possible and probable faults [50]. The exploration of faults can be supported by methods such as fault tree analysis (FTA), failure mode and effects analysis (FMEA) and benchmarking as well as by model-based requirements management (compare Holder et al. [20]).

The second level in Figure 2 is the most concrete level of product description and contains the product structure and geometry as well as the detailed material choice. The modular structure of the product is also situated on this level. A large body of research covers the placement of sensors and actuators; the results are methods and algorithms which ensure optimised geometrical placement [46]. The most prominent measure to increase fault-tolerance are redundant sensors or actuators. However, the application of concrete redundant elements can lead to considerable disadvantages in terms of cost, weight and required space [50]. Additionally, concrete redundant elements are sensitive to the same problems and faults. Consequently, it may happen that two or more redundant elements fail at the same time. It is therefore extremely important to include the more abstract levels into a holistic fault-tolerant design. Other approaches on the most concrete level are "over-actuation" (stronger or more actuators than necessary for the direct functionality; the excessive actuation potential can lead to better controllability and allow fault accommodation) and "overlap" (sensor overlap are zones which more than one sensor covers – the comparison of the sensor readings for the same area can be used for calibration purposes and for sensor fault detection). In general, the established concepts of "Design for Safety" (DfS) such as "Safe-Life" (compare e.g. [52]) and "Fail-Safe" lead to an increase of fault-tolerance; the approaches for "inherently safe design" can be adapted to fault-tolerant design (compare [50]).

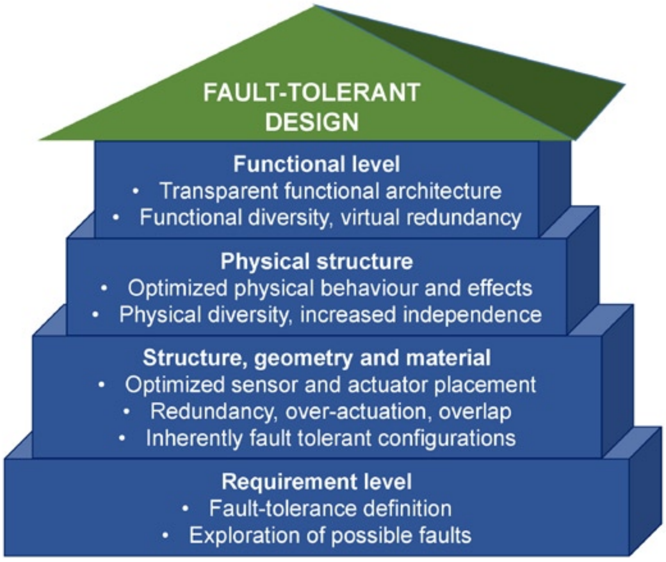


Fig. 2. Fault-tolerant design on different levels

The "physical structure" of a product describes the physical phenomena which realize the functionality of a product. A conscious design on this level can support the application of fault-tolerant control, e.g. when effects are used which can easily be monitored. For this purpose, methods and tools of design science can be adopted which support modelling and synthesis on this level, such as the use of demarcated physical effects (compare [17]). Very important on the level is the possibility to include physical diversity into redundant elements. For instance, if redundant sensors are used, it can be very advantageous to base them on different physical effects. In this case, certain faults such as extreme fog would only influence one of the sensors.

On the "functional level" of a product, the operations are described which take place in order to transform the state of an entity of the technical system into another state. For instance, the function that an electrical motor performs, is to transform electrical energy at the input connectors (input state) into mechanical energy at the output shaft (output state). Elaborate approaches to model the functions of technical systems were developed in the last years, e.g. the integrated function modelling framework (IFM – [17]) or the integration in an engineering framework based on graph-based design languages [58]. It is important to note that the highest and most independent form of redundancy can be achieved through diversity on the functional level. For instance, a physical sensor could be replaced by a virtual sensor based on an analytic redundancy. This sensor would work differently even on the function level.

On the different levels, fault-tolerant design characteristics were developed which were applied to a gear shifting system for a formula student driverless race car. These characteristics influence the reliability of the race car in multiple ways. In this competition, the success of a racing team is ultimately determined by its innovative capabilities and the skills to realize reliable solutions. A car with the potential to win the competition needs to be light-weight, energy-efficient and extremely reliable. Indeed, robustness and fault-tolerance is very important, as most of the evaluation points are directly or indirectly linked with the enduring of longer parts of the event. Also the maintainability is an important requirement, as the possibilities and time for maintenance are limited at the events of the competition. In order to meet these requirements, a gear shifting system was developed, which is explained in the next section

3. Automated shifting system

The gear box of the race cars of the Ravensburg-Weingarten University (RWU) is the integrated sequential gear box belonging to the four-cylinder in line motorcycle engine (Honda CBR 600), which is also used in the race-car – both the autonomous and the conventional combustion car. In a motor-bike, this gear-box is manually operated through a foot lever. Obviously, this is not possible for a driverless vehicle, therefore an automated shifting system is required. Initially the team developed a pneumatic shifting system, which is shown in Figure 3.

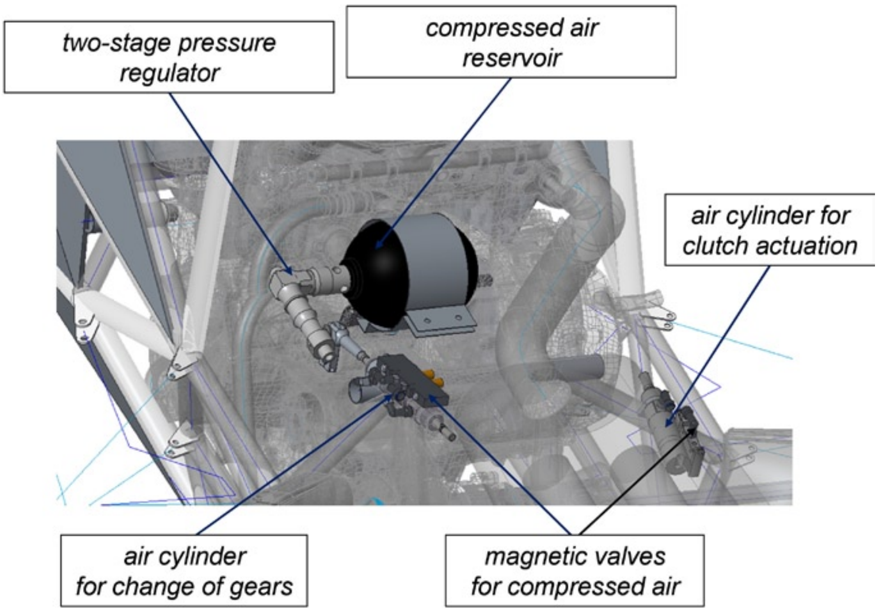


Fig. 3. Prior pneumatic gear shifting system

The pneumatic gear system consists of an air reservoir for compressed air, a two-stage pressure regulator, an air cylinder for gear changing, an air cylinder for clutch actuation and magnetic valves. Though this system was working in the race car, it still disposed of several disadvantages. The system consumed too much space, is relatively heavy and rather slow. With this gear shifting system shifting times of approx. 300ms can be achieve (for comparison: the later developed electromechanical system, which is described in the next section, is able to achieve 45ms). Because of the disadvantages in terms of space, weight and switching velocity, the team made the decision to develop a new gear shifting system. On the basis of literature reviews and benchmarking it was found out that an electromechanical solution has the potential to address all three disadvantages. The newly developed gear shifting system is shown in Figure 4.

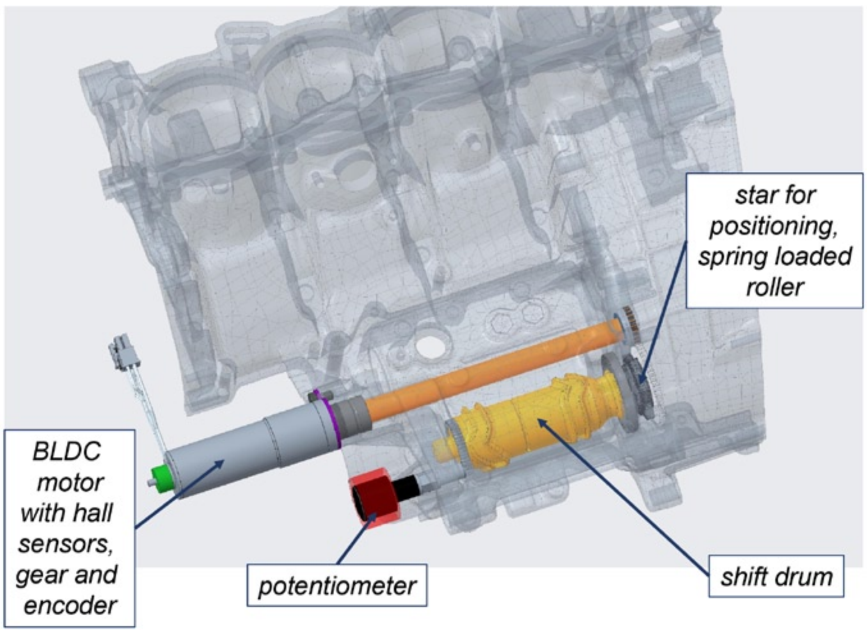


Fig. 4. Newly developed electromechanical gear shifting system

This gear shifting system centres on an electrical motor equipped with an incremental encoder. This incremental encoder allows obtaining the exact rotary position of the shaft of the electrical motor, if a predefined initialization run is carried out (e.g. for the first gear a position value of 22100 inc is assigned - this matches 138,7° at the gear exit and 60° at the shift drum). By means of a long shaft and a pair of spur gears the motion of the output shaft of the electrical motor is transferred to the modified shift drum. This shift drum causes a movement of the gear sleeves that allow the selection of a gear; this corresponds to the original situation in the motor-bike gear set. This shift drum also is connected to a potentiometer, located at the end opposite of the pair of spur gears. This potentiometer delivers an analogous resistance which is depending of the rotary position of the shifting drum.

In the next section, the design aspects which lead to an increase of the fault-tolerance of the gear shifting system are explained in detail.

4. Fault-tolerant design of the shifting system

This section discusses the design aspects which increase the fault-tolerance; the structure follows the levels of product concretization shown in Figure 2. The basis for any successful development is the requirement level, which describes the design objectives; design aspects on this level are described in Section 4.1. The most concrete level of product description contains the product structure and geometry as well as the detailed material choice. This level describes the technical system in detail; Section 4.2 describes design aspects on this level. The more abstract level "physical structure" describes the physical phenomena which realize the functionality of a product; design aspects on this level are contained in Section 4.3. Even more abstract is the "functional level" of a product, within which the operations are described which take place in order to transform the states of entities of the technical system into other states; Section 4.4 summarizes the design aspects on this level. A conscious fault-tolerant design takes place on all of these levels of product concretization and a link to fault-tolerant control is also possible and sensible on all these levels. For instance, over-actuation on the geometry level can enhance the possibilities of a fault-tolerant controller to accommodate faults and a virtual redundancy on the functional level can enable to fault detection and identification block of a fault-tolerant control scheme. The subsequent subsection will explain concrete measures of fault-tolerant design applied to the gear shifting system.

4.1. Design aspects on the requirements level

As stated above, the requirement level is extremely important for successful product development processes. In the given project, the initial step – search for requirements – contained a detailed analysis of the rules, which are each year published by the different formula student competitions; an important one is published by Formula Student Germany (FSG) and has a length of 133 pages. It is important to identify explicit requirements in this kind of large document and to make them traceable. Further requirements result from a conscious analysis of all possible driving, production and maintenance scenarios of the race-car. Such activities lead to a large number of requirements. This large number reduces the risk of failure caused by unknown or forgotten requirements, but requires a conscious management. Today, several software solutions for requirements management are available; Eclipse ProR is an open source option. Figure 5 shows requirements concerning the automated gear shifting system modelled in Eclipse ProR.

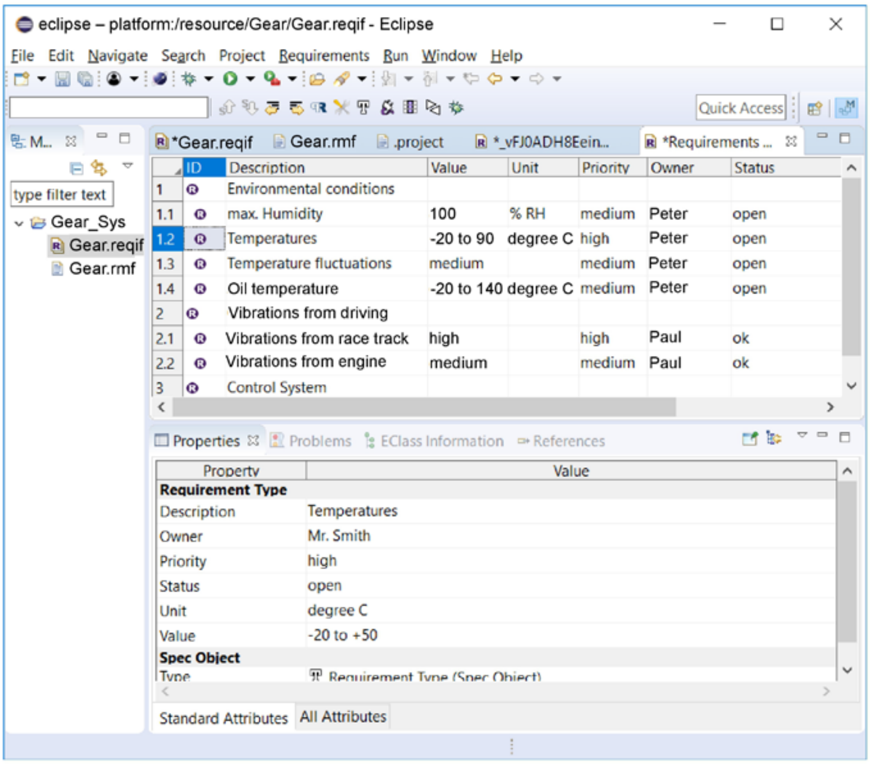


Fig. 5. Requirements of a electromechanical gear shifting system in Eclipse ProR

For fault-tolerant design, an identification of possible and probable faults needs to be performed. Similar to the collection of requirements, a conscious analysis of all possible driving, production and maintenance scenarios of the race-car can offer a basis for this activity. The most important identified faults for the gear shifting system were a tooth-on-tooth situation of the gear wheels, which will prohibit a gear shift and a loss of the position information of the shift drum.

4.2. Design aspects on the structure, geometry and material level

Obviously, it is possible to raise the fault-tolerance of the gear shifting system on the most concrete level "structure, geometry and material". In the current development, two cases of redundancy and the principle "over-actuation" could be integrated in the final design. The important elements of the gear shifting system are shown in Figure 6.

The electrical BLDC (brushless, direct current) motor (maxon EC-I 30), which is equipped with hall sensors, a gear system for increased torque (maxon GP 32 C) and an incremental encoder (maxon ENX 16 EASY), drives the gear shifting system. The controller of this motor is an electronic position control (maxon EPOS4). The output of the gear system is connected to a long shaft that transfers the torque to the other side of the gear system. The end of the shaft at this side is connected to a spur gear. This spur gear is contact with another spur gear, which is connected to the shift drum. At this point of connection, also a module is connected which consists of an element formed like a star and a roller which is pressed against this stare by means of a spring. This module is forcing the shift drum to certain, equally distributed rotary positions. Each of these position corresponds to a position of the shift drum within which a gear is exactly engaged. In this kind of gear systems, the shift drum disposes of geometrical entities (similar to a ditch) which are moving shift forks; those entities are visible in Figure 6. The shift forks connect the gear wheel of a given gear to one of the shafts of the gear systems, thus allowing a torque flow through the gear system. At the end of the shift drum another set of spur gears is located, which drive a potentiometer. Therefore, this potentiometer is able to monitor the angular position of the shift drum.

Obviously, this potentiometer is a second means to determine the angular position of the shift drum, because this information can also be deduced from the information given by the incremental encoder at the drive motor. This redundancy is the first aspect of fault-tolerant design that could be realised in the gear shifting system. It is important to note that these two means of measurement rely on totally different physical principles. Therefore, they can be understood as located on a higher level of fault-tolerant design and will be discussed in the following section.

The module, which consists of the star element, the roller and the spring and supports a precise and stable angular position of the shift drum, is also a redundant elements, as the angular position could

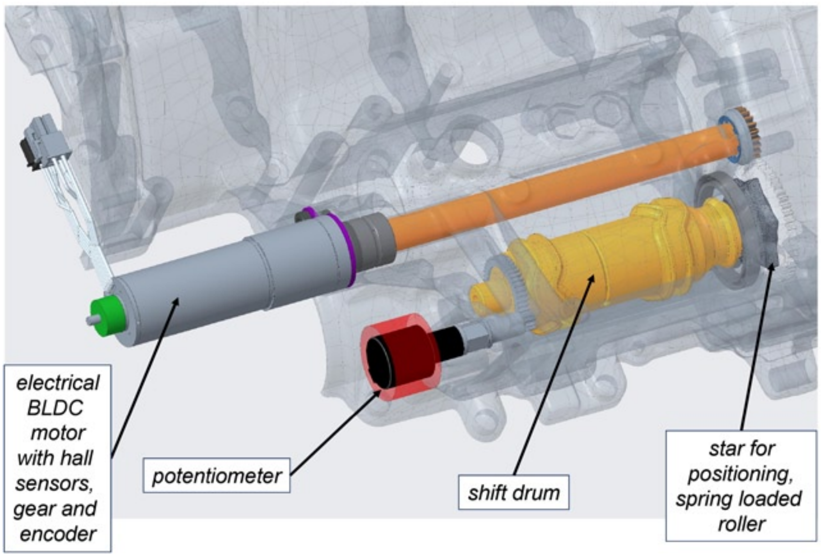


Fig. 6. Detail view of the shifting system

also be achieved by the electrical motor alone. In the original application of the gear system – the motor bike – the module is necessary for maintaining an optimum angular position for each gear. In the novel system, the motor is equipped with a self-looking gear (maxon GP 32 C – ration 14:1); theoretically this self-looking gear assures an exact positioning after a gear was properly engaged by the controlled motor. Still, during drive testing it became obvious that the control quality and gear engagement speed are both not affected by this redundant element and that the engaged gear is held even more stable and safely. Furthermore, the module supports the teaching of the position of the shift drum during the initial run, because it achieves optimum angular positions for each gear mechanically. Additionally, energy can be saved because of this module, since it allows to energize the shifting motor only temporarily.

The notion "over-actuation" can have two meanings [50]:

- on the one hand, it can indicate the employment of more actuators than ultimately necessary for realizing the given functionality and
- on the other hand, it can indicate the employment of stronger actuators than ultimately necessary for realizing the given functionality.

The main advantages of over-actuated systems are both an improved controllability and the capability to accommodate faults. This capability is an effect of the fact that the over-actuation potential can be employed for compensating the consequences of one or more faults [50]. Consequently, over-actuation can be a means to increase the fault-tolerance of a technical system. In the novel gear shifting system, over-actuation could be realised in the electrical motor. Initially, the necessary torque for turning the system of shaft and shift drum was found be using a wrench for actuation and measuring the applied torque. It became obvious that about 1.5 Nm are necessary for this turning operation. The chosen electrical motor with the gear system is able to generate a torque of up to 6 Nm (only for a short time; however, this time is in any case long enough for one or more gear engagement processes). The later testing lead to the insight that this over-actuation leads to fast shifting times and also a superior control quality.

4.3. Design aspects on the physical level

The physical level connects the concrete solutions in terms of structure, geometry and material with the abstract solution description on the function level. Since several years, a conscious development of the physical level is advised in design science (e.g. [15]), because it allows innovative solutions. During an analysis of so-called "break-through products", i.e. product which are innovative to such an extent that their successor products are immediately outdated, lead to the observation that nearly all break-through products dispose of altered physical phenomena for realising the central functionality. Additionally, an in-depth occupation with the physical phenomena which realize the numerous functions of a technical system can support a better system understanding. Current research activities intend to expand the analysis of physical phenomena by means of including uncertainties in form of disturbances [34] and to integrate representations of physical phenomena in a holistic engineering framework. The outcome of an analysis of the physical phenomena (or "effects") is a chain of physical effects; Figure 7 shows an example of an incremental encoder, as used in the given project.

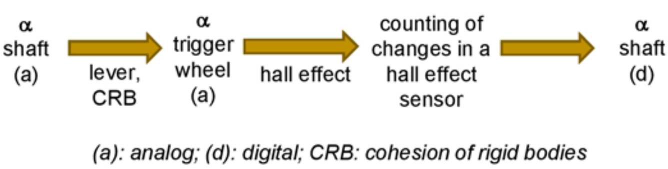


Fig. 7. Physical effect chain incremental encoder

The physical effects that enable the function of the incremental encoder are depicted in this effect chain. The angle of a shaft is transferred by means or the two physical effects "lever" and "cohesion of rigid bodies (CRB) to a trigger wheel. The Hall Effect and the counting of changes in a hall effect sensor leads to a digital information concerning the angle of the shaft. The experiences in the underlying project lead to the insight that the analysis of physical phenomena leads to a deepening of the understanding and assists communication processes.

A central possibility on this level to increase fault-tolerance is the design principle "physical diversity". This principle describes the conscious application of system entities that employ different physical phenomena for the achievement of given objectives [50]. Examples for physical diversity are given in aircraft industry, where an increase level of fault-tolerance is achieved by employing both hydraulic power lines with hydraulic actuators and electric power lines with electric actuators [7]. In the last section, one example of physical diversity was already initially described: the application of two sensors with one identical goal – the monitoring of the angular position of the shift drum. This monitoring result is extremely important for enabling a position control of the motor and for assuring that the gears are safely engaged.

As described above, the electrical motor is equipped with a gear system (transmission ratio 14:1). This enables the incremental encoder to distinguish 57.344 angular positions (1024 (pulses per channel per revolution) * 4 (resolution) * 14 (gear reduction)) of the shift drum. The information delivered by the motor position control unit (EPOS 4) are shown in Figure 8. It is important to note that this information are on the one hand resulting from the incremental encoder, on the other from the hall sensors in the motor, which are necessary for the control the brushless motor.

In Figure 8 upper part, the position demand and two kinds of actual positions (case without fault and case with fault) are shown. Similarly, the middle part shows the velocity demand and two kinds of actual velocity (case without fault and case with fault). The lower part concerns the current. For both the nominal case and the faulty case the actual current and the average actual current are shown. The essential content of Figure 8 is the investigation of an error during the engagement of the first gear (corresponding to an encoder reading of 22100 inc) from neutral (corresponding to an encoder reading of 0 inc). As can be seen in Figure 8, a certain amount of redundancy is already present because of the combination of electrical motor, incremental encoder and motor position control. The position control is able to determine a motor angular position; for this the information from the encoder is used. It is additionally able to determine the velocity of the motor at certain instances of time, by means of combined information from the encoder applying numerical differentiation and from the hall sensors of the motor (the presence of the hall sensors in the motor is also a form of redundancy). Furthermore, the position control is able to determine the current, which it is delivering to the motor when trying to achieve an intended angular position. This current is a result of the PID controller implemented within the motor control as a low level control loop.

In the developed design, a further sensor was included – the potentiometer mentioned above – that also has the intention to monitor the angular position of the shift drum. The measurable resistance of this potentiometer is not assessed by the position control but by the superordinate control unit of the race-car. As a consequence of this constellation, in superordinate control unit the engaged gear will be known, even in the case of a fault concerning the electrical motor or the motor position control. In such cases, the race-care would be able to finish the race in the gear which was engaged when the fault occurred. The superordinate control unit could open the clutch for speeds which are too low for this engaged gear (in such cases the mo-

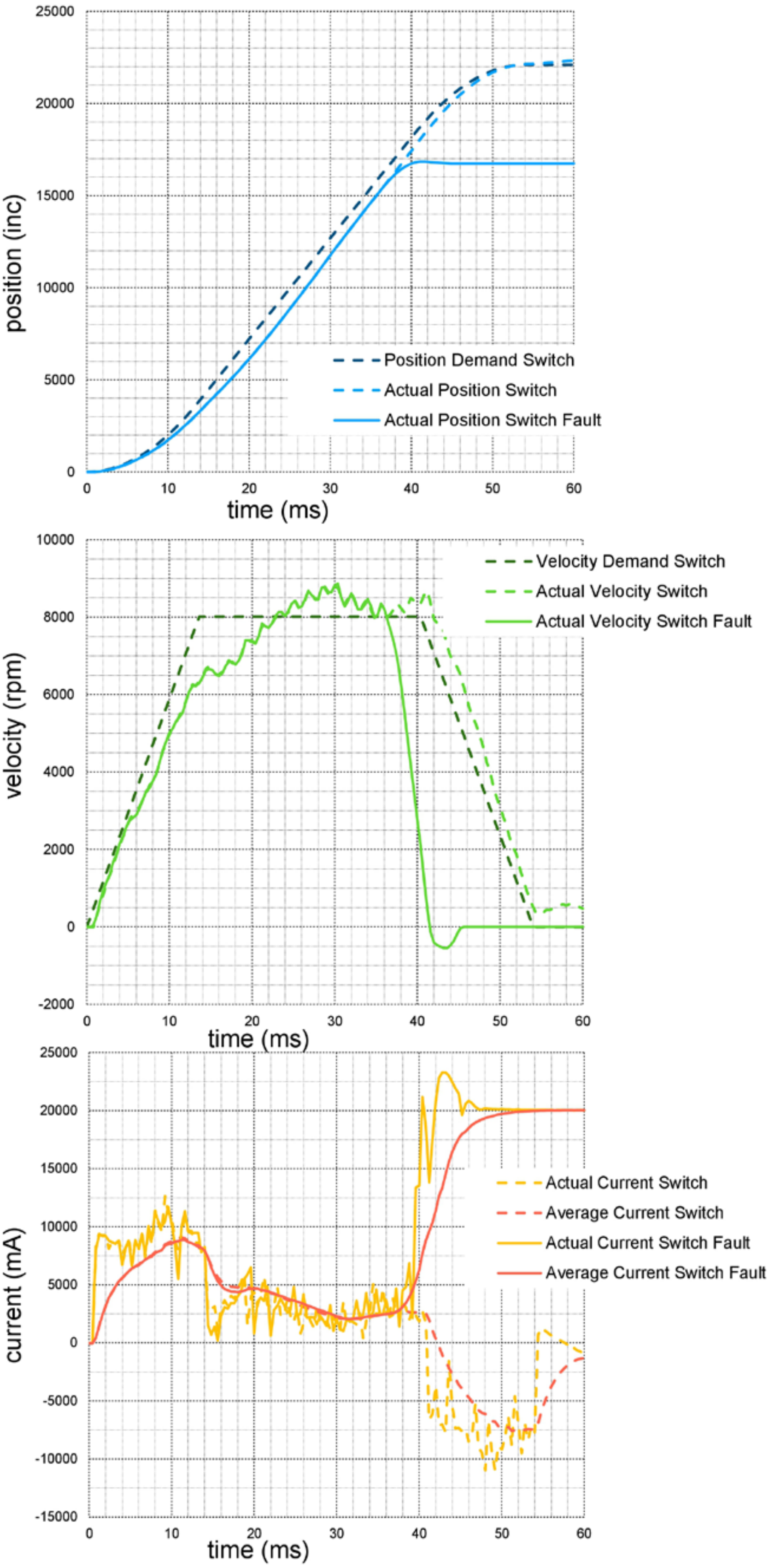


Fig. 8. Position over time (top), velocity over time (middle) and current over time (bottom)

tor could be forced into very low speeds and could stall, if the clutch is not opened). Even a standstill of the car would be possible in this situation, because with an adapted clutch characteristic the race-car can start form standstill also in the 2nd and 3rd gear (most competitions are carried out on nearly even race tracks; this allows start in higher gears). As mentioned above, it is of paramount importance to finish a course in case of a fault, because otherwise no points are awarded. This example shows clearly, which advantages in a competition can be achieved with an increased fault-tolerance leading to an increase reliability. It is important to note that a sensor redundancy enables the application of sensor fusion techniques. These techniques apply algorithms such as the least squares method or the Kalman filter, which can lead to higher accuracy and credibility [2]. In the realised system, up to now, only a plausibility check was applied; sensor fusion techniques will be incorporated in the next development generation.

4.4. Design aspects on the function level

In the last years new powerful tools were developed for function modelling, most notably the integrated function modelling framework (IFM – [17]), which can also be integrated in current engineering frameworks [19]. Figure 9 shows a function model of the gear shifting system modelled in IFM.

In this modelling framework, information is presented in associated views. The state view (upper left part of the IFM) represents states of actors and operands as well as their change. The process flow view (upper right part of the IFM) represents the flow of transformation and interaction processes. This two views are accompanied by an interaction view and an actor view (lower part of the IFM). This framework fosters a holistic view on the functional level. For instance, in the IFM redundancy of incremental encoder and potentiometer is clearly visible. Two main operands "energy" and "signal" are modelled; their flow through the system is also obvious in the IFM.

As mentioned in Section 4.1, one prominent fault is a "tooth-on-tooth" situation, when a gear should be engaged; this fault is a main focus of the design aspects for increased fault-tolerance on this level. The position control of the electrical motor can deliver crucial information for the accommodation of this fault; these information are shown in Figure 10. It is possible to use these information together with an analytical model for the detection of this fault and its accommodation.

Essentially this fault appears, when the side surfaces of the gear that should be moved to engage a gear are aligned and consequently do not allow this engaging movement, which is initiated by the shift forks. When looking at the information presented in Figure 10, it is apparent that this fault can be detected, if the both the velocity over position and the current over position reach certain regions (visible in red hatched in Figure 10). Thus, the red hatched zones represent the fault signature of this fault. In the case that both redundant information are present (velocity over position and current over position), it is clear that the fault "tooth-on-tooth" is present. In the given case it was satisfactory to detect certain regions in the velocity – position diagram and the current – position diagram, i.e. certain combinations of these parameters which indicate this fault. In more complex cases other algorithms such as the well-known observer-based fault detection method could allow the detection and identification of the fault [18].

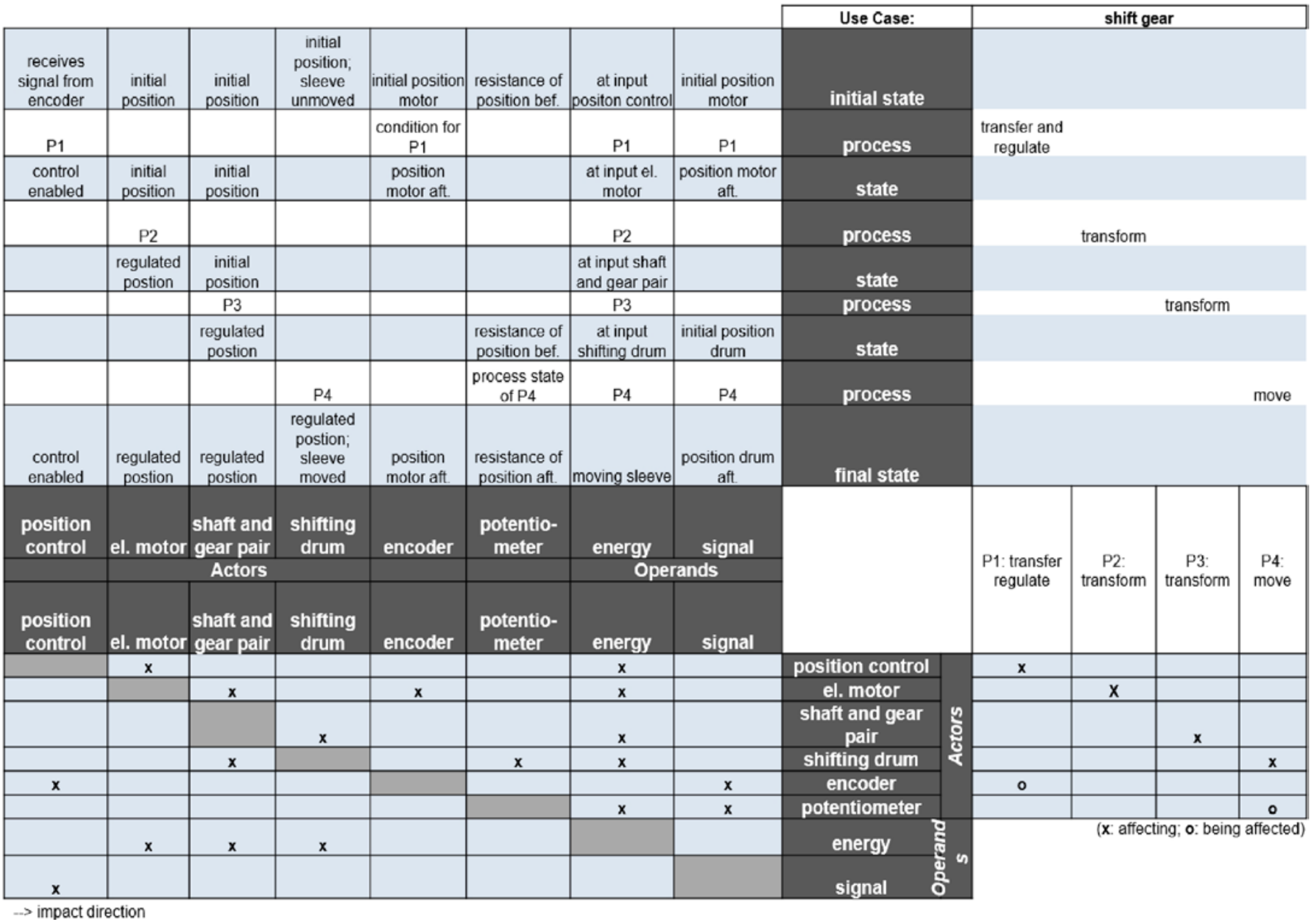


Fig. 9. IFM model of the gear shifting system

For accommodating this fault it is possible to let the electrical motor for gear shifting reverse for a few degree and then to make another attempt to move the shift drum, which moves the shift fork, which may engage the gear. The testing showed that, in a large share of the cases when this fault occurred, this simple manoeuvre was sufficient to enable the engagement of the gear. In these cases, only a small amount of time is lost (10 to 20 ms) and the shifting is still very fast. A combined possible measure for the accommodation of this fault is a short engagement of the clutch (induced by the superordinate control system of the race car), which can realize a small rotation of the gears which may resolve the "tooth-on-tooth" situation.

Furthermore, the elaborate control and diagnosis system can also lead to an improved overall performance of the race-car as a consequence of further decreased shifting times. In the earlier cars, the ignition was cut-off during the whole shifting process until the gear was definitely engaged (this ignition cut-off is needed to have smaller contact forces on the gears and to allow shifting). In the new system, it is possible to be definitely sure that a gear will be engaged, even before the final position is achieved, because of the integrated monitoring of position and current. In this case, the superordinate control may reconnect the ignition and increase the injection quantity quite a bit earlier. In the moment, when the gear is finally engaged, the engine will again be able to deliver its full torque, leading to optimised performance. In such cases, the superordinate control will overrule the gear-ignition-cut control.

6. Conclusions and outlook

The international competition "formula student" has gained enormous attention over the last years. The demands concerning performance and reliability are immense. As a consequence of the inclusion of the capability to drive autonomous, the complexity of the cars and especially the control systems has increased dramatically. This leads to additionally fault possibilities and the importance of fault-tolerance has also increased immensely. The same is true for complex technical systems produced in industrial companies. The research activities concerning fault-tolerant control address these issues since several years; in the last years also a focus on fault-tolerant design is visible. Fault-tolerant design pursues two main goals:

- … to allow and ease effective fault-tolerant control by means of conscious design aspects.
- … to integrate design aspects which by themselves increase the fault-tolerance (e.g. inherently fault-tolerant design aspects).

The focus of this paper is an in-depth discussion of the methods and tools of fault-tolerant design. This discussion as based on the conscious development of fault-tolerant design aspect for an automated shifting system for a formula student race car. A model, which allows a distinction of the level of abstraction of the models of the technical system under development, served for structuring this discussion. On all levels, concrete measures are explained that improve the fault-tolerance of the race-care and, through this, its reliability. Some of the measures had the additional effect to optimise the performance of the race-car. Interestingly, the implemented measures only lead to the addition of one element (a potentiometer) and they did not lead to nega-

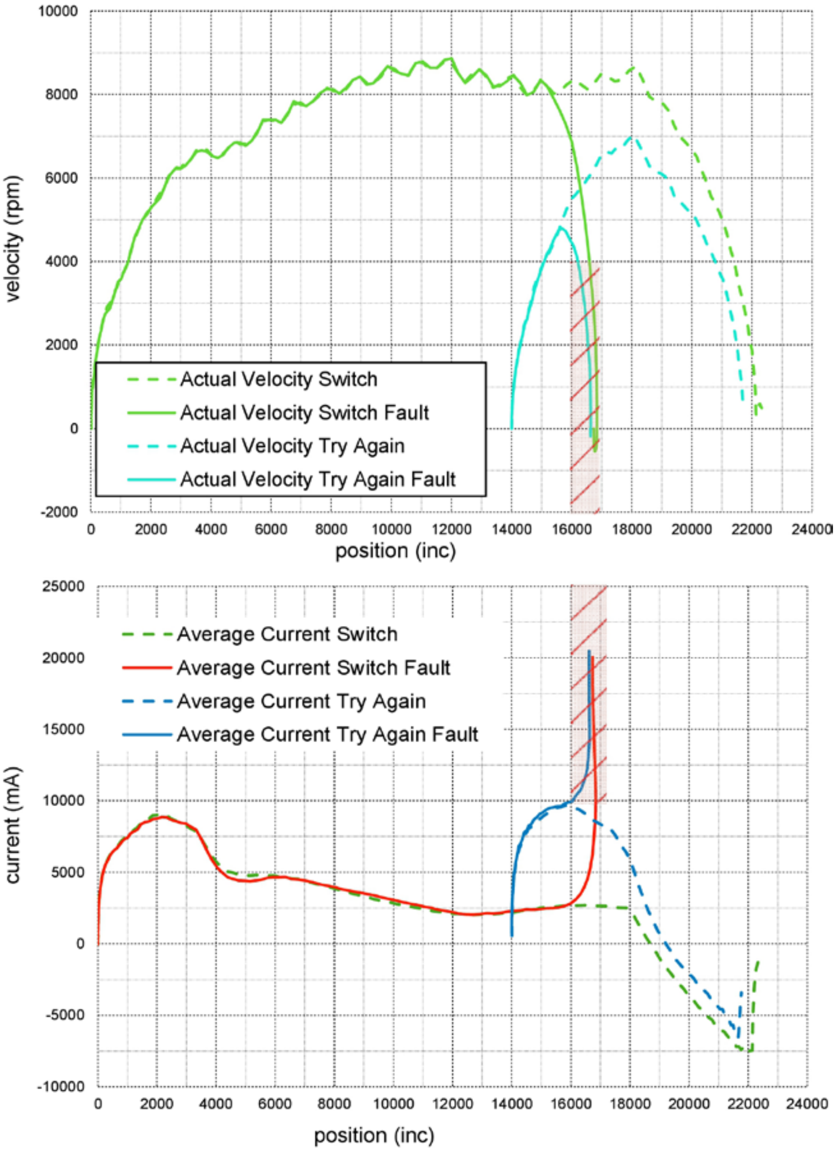


Fig. 10. Velocity over position (top) and current over position (bottom)

tive consequences in terms of space, weight and money. It is important to note that the measures are based on an intelligent mechanical and electromechanical design in combination with an elaborate diagnosis and control systems. It was possible to identify a positive influence resulting from the application of the methods and tools of design science, for instance integrated function modelling and a conscious analysis of physical phenomena. Future research activities in this promising field need to expand the methodical basis of fault-tolerant design. These activities need to include technical systems with different sets of requirements, application scenarios and levels of complexity.

Acknowledgement

The work was partially supported by the National Science Centre, Poland under Grant: UMO-2017/27/B/ST7/00620 and in the scope of the project „digital product life-cycle (ZaFH)“, which is supported by a grant from the European Regional Development Fund and the Ministry of Science, Research and the Arts of Baden-Württemberg, Germany (https://efre-bw.de/).

References

1. An H, Fidan B, Liu J, Wang C, Wu L. Adaptive fault-tolerant control of air-breathing hypersonic vehicles robust to input nonlinearities. International Journal of Control 2019; 92(5): 1044-1060, https://doi.org/10.1080/00207179.2017.1381346.
2. Banachowicz A, Wolski A. A Comparison of the Least Squares with Kalman Filter Methods Used in Algorithms of Fusion with Dead Reckoning Navigation Data. Transnav, the International Journal on Marine Navigation and Safety of Sea Transportation 2017; 11, 4; 691 – 695, https://doi.org/ 10.12716/1001.11.04.16.
3. Barbieri G, Fantuzzi C. Borsari R. A model-based design methodology for the development of mechatronic systems. Mechatronics 2014; 24: 833 – 843, https://doi.org/10.1016/j.mechatronics.2013.12.004
4. Bernard R, Irlinger R. About watches and cars: winning R&D strategies in two branches. Presentation at the International Symposium "Engineering Design – The Art of Building Networks"; Garching, Germany, 2016.
5. Bianchi N, Dai Pré M, Bolognani M. Design of a Fault-Tolerant IPM Motor for Electric Power Steering, IEEE Transactions on Vehicular Technology 2006; 55 (4): 1102 – 1111, https://doi.org/10.1109/TVT.2006.877716.
6. Blanke M, Kinnaert M, Lunze J, Staroswiecki M. Diagnosis and Fault-Tolerant Control. New York: Springer, 2016, https://doi.org/10.1007/978-3-662-47943-8.
7. Brando G, Dannier A, Del Pizzo A, Di Noia L P. Electric steering for aircraft nose landing gears using axial-flux permanent-magnet motors. Proceedings of the XXII International Conference on Electrical Machines (ICEM) 2016, https://doi.org/10.1109/ICELMACH.2016.7732612.
8. Chamas M, Paetzold K. Modeling of Requirement-Based Effect Chains of Mechatronic Systems in Conceptual Stage. International Journal of Electrical and Electronic Engineering & Telecommunications 2018; 7 (3): 127-134, https://doi.org/10.18178/ijeetc.7.3.127-134.
9. Chen S, Chen B, Shi F. Distributed Fault-Tolerant Consensus Protocol for Fuzzy Multi-Agent Systems. Circuits Systems Signal Process (2019) 38:611–624, https://doi.org/10.1007/s00034-018-0872-y.

10. Charrier J J, Kulshreshtha A. An electric actuation for flight and engine control system: Evolution, current trends and future challenges. In Proceedings of the 45th AIAA Aerospace Sciences Meeting and Exhibit, 2007, https://doi.org/10.2514/6.2007-1391.
11. Cordoneanu D, Nitu C. A Review of Fault Diagnosis in Mechatronics Systems. In: Gheorghe G. (eds) Proceedings of the International Conference of Mechatronics and Cyber-MixMechatronics – 2018. ICOMECYME 2018. Lecture Notes in Networks and Systems, vol 48. Springer, Cham, https://doi.org/10.1007/978-3-319-96358-7_18.
12. Cross N. Engineering Design Methods: Strategies for Product Design. John Wiley and Sons Ltd., 2008.
13. Cui J, Ren Y, Xu B, Yang D, Zeng S. Reliability analysis of a multi-eso based control strategy for level adjustment control system of quadruped robot under disturbances and failures. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 42–51, http://dx.doi.org/10.17531/ein.2020.1.6.
14. Dorociak R, Gausemeier J. Modeling of the Failure Propagation of an Advanced Mechatronic System within the Specification of its Principle Solution. Proceedings of the Design Society: Design Conference 2012: 807– 816.
15. Dubrova E. Fault-Tolerant Design. Springer, 2013. https://doi.org/10.1007/978-1-4614-2113-9.
16. Ehrlenspiel K, Meerkamm H. Integrierte Produktentwicklung. Denkabläufe, Methodeneinsatz, Zusammenarbeit. Carl Hanser, 2013, https://doi.org/10.3139/9783446436275.
17. Eisenbart B, Gericke K, Blessing L, McAloone T. A DSM-based Framework for Integrated Function Modeling: Concept, Application and Evaluation. Research in Engineering Design, 2016; 28 (1): 25 - 51. https://doi.org/10.1007/s00163-016-0228-1.
18. Eissa M A, Darwish R R, Bassiuny A M. Design of Observer-Based Fault Detection Structure for Unknown Systems using Input–Output Measurements: Practical Application to BLDC Drive. Power Electronics and Drives 2019: 4, 39; 217 – 226, https://doi.org/ 10.2478/pead-2019-0017.
19. Elwert M, Ramsaier M, Eisenbart B, Stetter R. Holistic Digital Function Modelling with Graph-Based Design Languages. In: Proceedings of the Design Society: International Conference on Engineering Design / Volume 1 / Issue 1, Cambridge University Press, 2019: 1523-1532, https://doi.org/10.1017/dsi.2019.158.
20. Gausemeier J, Pöschl M, Dyetr S, Kaiser L. Modeling and Analyzing Fault-tolerant Mechatronic Systems. Proceedings of the Design Society: International Conference on Engineering Design ICED 2009: 6-55 – 6-66.
21. Gao Z. Active Disturbance Rejection Control: A Paradigm Shift in Feedback Control System Design. Proceedings of the 2006 American Control Conference. IEEE: 2006.
22. Han J. From PID to Active Disturbance Rejection Control. IEEE Transactions on Industrial Electronics 2009; 56 (3): 900 – 906, https://doi.org/10.1109/TIE.2008.2011621.
23. Holder K, Zech A, Ramsaier M, Stetter R, Niedermeier H-P, Rudolph S, Till M. Model-Based Requirements Management in Gear Systems Design Based On Graph-Based Design Languages. Applied Sciences, 2017; 7, 1112; https://doi.org/10.3390/app7111112.
24. Huang H-Z, Yu K, Huang T, Li H, Qian H-M. Reliability estimation for momentum wheel bearings considering frictional heat. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 6–14, http://dx.doi.org/10.17531/ein.2020.1.2.
25. Hruschka P. Business Analysis und Requirements Engineering: Produkte und Prozesse nachhaltig verbessern. Hanser, 2014, https://doi.org/10.3139/9783446438620.
26. Hsieh T-Y, Li K-H, Chung C-C. A fault-analysis oriented re-design and cost-effectiveness evaluation methodology for error tolerant applications. Microelectronics Journal, 2017; 66: 48 – 57, https://doi.org/10.1016/j.mejo.2017.05.018.
27. Iscioglu F, Kocak A. Dynamic reliability analysis of a multi-state manufacturing system. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2019; 21 (3): 451–459, http://dx.doi.org/10.17531/ein.2019.3.11.
28. Isermann R. Fault Diagnosis Systems. An Introduction from Fault Detection to Fault Tolerance. New York: Springer, 2006, https://doi.org/10.1007/3-540-30368-5.
29. Li J, Wang Z, Ren Y, Yang D, Lv X. A novel reliability estimation method of multi-state system based on structure learning algorithm. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 170–178, http://dx.doi.org/10.17531/ein.2020.1.20.
30. Li Y, Wang K. Modified convolutional neural network with global average pooling for intelligent fault diagnosis of industrial gearbox. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 63–72, http://dx.doi.org/10.17531/ein.2020.1.8.
31. Lin J W, Yang, M F. Fault-tolerant design for wide-area mobile ipv6-networks. The Journal of Systems and Software, 2009; 82, 1434 – 1446, https://doi.org/10.1016/j.jss.2008.07.021.
32. Lindemann U. Methodische Entwicklung technischer Produkte. Springer, 2009, https://doi.org/10.1007/978-3-642-01423-9.
33. Konowrocki R, Chojnacki A. Analysis of rail vehicles' operational reliability in the aspect of safety against derailment based on various methods of determining the assessment criterion. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 73–85, http://dx.doi.org/10.17531/ein.2020.1.9.
34. Mathias J, Eifler T, Engelhardt R, Kloberdanz H, Birkhofer H, Bohn A. Selection of Physical Effects Based on Disturbances and Robustness Ratios in the Early Phases of Robust Design. In: Proceedings of the International Conference on Engineering Design, ICED11, 15 - 18 August 2011, Technical University of Denmark.
35. Mazurkiewicz D. Computer-aided maintenance and reliability management systems for conveyor belts. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (3): 377–382.
36. Mhenni F, Nguyen N, Choley J Y, Rivière A. SafeSysE: A Safety Analysis Integration in Systems Engineering Approach, IEEE Systems Journal 2018; 12(1) 161 – 172, https://doi.org/10.1109/JSYST.2016.2547460.
37. Mu L, Li L, Yu X, Zhang Y, Li P, Wang X. Observer-based fault-tolerant control of hypersonic scramjet vehicles in the presence of actuator faults and saturation. International Journal of robust and nonlinear control 2019; 29(16): 5377-5393, https://doi.org/10.1002/rnc.4004.
38. Münzer C, Shea K. Simulation-Based Computational Design Synthesis Using Automated Generation of Simulation Models from Concept Model Graphs. Journal of Mechanical Design 2017; 139(7), 071101, https://doi.org/10.1115/1.4036567.
39. Nie S, Mao C, Wang D. Fault Tolerant Design for Electronic Power Transformer. In Proceedings of the IEEE PES Asia-Pacific Power and Energy Conference 2016, https://doi.org/10.1109/APPEEC.2016.7779592
40. Njindam T, Paetzold K. Design for Reliability: an Event- and Function-Based Framework for Failure Behavior Analysis in the Conceptual Design of Cognitive Products. In: Proceedings of the Design Society: International Conference on Engineering Design ICED 2011.

41. Oh Y G, Jeong J K, Lee J J, Lee Y H, Baek S M, Lee S J. Fault-tolerant design for advanced diverse protection system. Nuclear Engineering and Technology 2013; 45(6): 795 – 802, https://doi.org/10.5516/net.02.2013.526.
42. Pahl G, Beitz W, Feldhusen J, Grote K-H. Engineering Design: a systematic Approach. Springer, 2007, https://doi.org/10.1007/978-1-84628-319-2.
43. Pham H. System Software Reliability. Springer, 2006, https://doi.org/10.1007/1-84628-295-0.
44. Ponn J, Lindemann U. Konzeptentwicklung und Gestaltung technischer Produkte. Springer, 2011, https://doi.org/10.1007/978-3-540-68563-0.
45. Ramirez-Neria M, Sira-Ramirez H, Garrido-Moctezuma R, Luviano-Juarez A. Linear active disturbance rejection control of underactuated systems: The case of the Furuta pendulum. ISA Transactions 2014; 53: 920 – 928, http://dx.doi.org/10.1016/j.isatra.2013.09.023.
46. Rathi S, Gupta R. Sensor placement methods for contamination detection in water distribution networks: a review. Procedia Eng. 2014; 89: 181 – 188, https://doi.org/10.1016/j.proeng.2014.11.175.
47. Rouissi F, Hoblos G. Fault tolerant sensor network design with respect to diagnosability properties. In: Proceedings of the 8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (SAFEPROCESS), 2012: 1120 – 1124, https://doi.org/10.3182/20120829-3-mx-2028.00067.
48. Saeed M, Rodriguez A, Arias M, Briz F. Flexible and Fault Tolerant Distributed Control Structures for Modular Power Electronic Transformers. Proceedings of the IEEE Applied Power Electronics Conference and Exposition (APEC), IEEE Xplore 2019, https://doi.org/10.1109/APEC.2019.8722316.
49. Shirazipourazad S, Sen A, Bandyopadhyay S. Fault-tolerant design of wireless sensor networks with directional antennas. Pervasive and Mobile Computing 2014; 13: 258 – 271, https://doi.org/10.1016/j.pmcj.2014.03.004.
50. Stetter R. Fault-Tolerant Design and Control of Automated Vehicles and Processes. Insights for the Synthesis of Intelligent Systems. Cham: Springer, 2020, https://doi.org/10.1007/978-3-030-12846-3.
51. Stetter R, Witczak M, Pazera M. Virtual Diagnostic Sensors Design for an Automated Guided Vehicle. Applied Sciences, 2018, 8(5), 702; https://doi.org/10.3390/app8050702.
52. Tavares S M O, de Castro P M S T. Damage Tolerance of Metallic Aircraft Structures Cham: Springer, 2019, https://doi.org/1010.1007/978-3-319-70190-5.
53. Tian J, Zhou C, Yang Y, Wu W, Mao C, Wang D. Individual DC Voltage Balance Control for Cascaded H-Bridge Electronic Power Transformer With Separated DC-Link Topology. IEEE Access 2019; 7: 38558 – 38567, https://doi.org/ 10.1109/ACCESS.2019.2905006.
54. 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.
55. Valis D, Pietrucha-Urbanik K. Utilization of diffusion processes and fuzzy logic for vulnerability assessment. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (1): 48–55.
56. Varga A. Solving Fault diagnosis Problems. Linear Synthesis Techniques. Springer, 2017, https://doi.org/10.1007/978-3-319-51559-5.
57. Vedachalam N, Umapathy A, Ramadass G A. Fault-tolerant design approach for reliable offshore multi-megawatt variable frequency converters. Journal of Ocean Engineering and Science 2016; 1: 226 – 237, https://doi.org/10.1016/j.joes.2016.06.001.
58. Vogel S. An application-independent continuum mechanics interface for virtual engineering. Engineering with Computers 2019; 35: 551 – 565. https://doi.org/10.1007/s00366-018-0617-3.
59. Witczak M, Majdzik P, Stetter R, Bocewicz B. Interval max-plus fault-tolerant control under resource conflicts and redundancies: application to the seat assembly. International Journal of Control 2019: 1 – 13; https://doi.org/10.1080/00207179.2019.1630749.
60. Zhang C, Jaimoukha I M, Sevilla F R S. Fault-tolerant observer design with a tolerance measure for systems with sensor failures. Proceedings of the American Control Conference (ACC) 2016, https://doi.org/10.1109/ACC.2016.7526861.
61. Zhang C, Zhang Y. Common cause and load-sharing failures-based reliability analysis for parallel systems. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 26–34, http://dx.doi.org/10.17531/ein.2020.1.4.
62. Zhang X, Zhao J. Compound fault detection in gearbox based on time synchronous resample and adaptive variational mode decomposition. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (1): 161 – 169, http://dx.doi.org/10.17531/ein.2020.1.19.
63. Zheng C, Hehenberger P, Le Duigou J, Bricogne M, Eynard B. Multidisciplinary design methodology for mechatronic systems based on interface model. Research in Engineering Design 2017; 28, 333 – 356. https://doi.org/10.1007/s00163-016-0243-2.

Ralf STETTER
Richy GÖSER
Sebastian GRESSER
Markus TILL
Department of Mechanical Engineering
Ravensburg-Weingarten University (RWU)
Doggenriedstrasse, 88250 Weingarten, Germany

Marcin WITCZAK
Institute of Control and Computational Engineering
University of Zielona Góra (UZ)
ul. Podgórna 50, 65-246 Zielona Góra, Poland

E-mails: ralf.stetter@rwu.de, rg-152598@hs-weingarten.de, sebastian.gresser@rwu.de, markus.till@rwu.de, m.witczak@issi.uz.zgora.pl

Article citation info:
ŠKERLIČ S, SOKOLOVSKIJ E, ERČULJ V. Maintenance of heavy trucks: an international study on truck drivers. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 493–500, http://dx.doi.org/10.17531/ein.2020.3.12.

Sebastjan ŠKERLIČ
Edgar SOKOLOVSKIJ
Vanja ERČULJ

MAINTENANCE OF HEAVY TRUCKS: AN INTERNATIONAL STUDY ON TRUCK DRIVERS

OBSŁUGA TECHNICZNA SAMOCHODÓW CIĘŻAROWYCH: MIĘDZYNARODOWE BADANIE KIEROWCÓW CIĘŻARÓWEK

Since the implementation of modern approaches in operation maintenance, drivers are expected to be integrated into the entire system of maintenance in order to take over the professional competencies of maintenance workers. For this purpose, an international study was conducted on a sample of 249 truck drivers, with the aim to determine how maintenance in transport companies affects the role of heavy truck drivers in fleet maintenance. Based on the developed SEM model, it was determined that building an efficient maintenance infrastructure in transport companies enables the active participation of truck drivers in the company's maintenance system. Drivers are capable of repairing minor failures, but only if they are proactive, which will affect their preventive behaviour. This can greatly benefit transport companies, as the results show that the vehicles are utilised more efficiently and in a better technical condition. The results therefore represent an important guideline for improving maintenance in transport companies, developing new competencies of truck drivers and upgrading existing knowledge in the development of modern operation maintenance.

Keywords: *maintenance, maintenance strategy, heavy trucks, truck drivers.*

Od czasu wdrożenia nowoczesnych metod obsługi technicznej, kierowcy powinni zostać zintegrowani z całym systemem obsługi technicznej, aby przejąć kompetencje zawodowe pracowników obsługi technicznej. W tym celu przeprowadzono międzynarodowe badanie na próbie 249 kierowców ciężarówek w celu ustalenia, w jaki sposób obsługa techniczna w firmach transportowych wpływa na rolę kierowców ciężarówek w utrzymaniu floty. Na podstawie opracowanego modelu równania strukturalnego (SEM) ustalono, że zbudowanie efektywnej infrastruktury serwisowej w firmach transportowych umożliwia aktywny udział kierowców ciężarówek w systemie obsługi technicznej firmy. Kierowcy są zdolne do naprawy drobnych usterek, ale tylko wtedy, gdy są proaktywni, co wpłynie na ich zachowanie zapobiegawcze. Może to być bardzo korzystne dla firm transportowych, ponieważ wyniki pokazują, że pojazdy są użytkowane bardziej wydajnie i w lepszym stanie technicznym. Wyniki stanowią zatem ważną wskazówkę dotyczącą poprawy obsługi technicznej w firmach transportowych poprzez rozwój nowych kompetencji kierowców ciężarówek i poszerzanie istniejącej wiedzy w zakresie rozwoju nowoczesnej obsługi technicznej.

Słowa kluczowe: *obsługa techniczna, strategia obsługi technicznej, samochody ciężarowe, kierowcy ciężarówek.*

1. Introduction

The introduction of modern maintenance strategies highlights a changed understanding of the organisation of maintenance, which leads to greater collaboration and teamwork in the company-wide maintenance process. An important feature of modern strategies is the introduction of independent machine and device maintenance by the operators, that is, connecting many basic service actions with the manufacturing process. Modern maintenance strategies therefore focus on two aspects of enhancing the competencies of employees: the first is to improve the competence of the staff responsible for operation maintenance. Another aspect is the inclusion of operators in the works for operation maintenance and the transfer of responsibility to them allows for a better utilisation of the knowledge they possess, reinforces their sense of their own value and makes them aware of their participation in the achievement of the enterprise's objectives [24].

However, there is a fundamental difference between the maintenance of machinery in manufacturing and fleet maintenance in companies. While researchers and maintenance engineers in manufacturing companies are thoroughly invested in enhancing the advanced actions of operators in machine maintenance [3], the maintenance of heavy truck drivers remains a neglected and under-researched field. Working in the transport industry has its own unique features, since the average driver in international transport spends five days a week in a truck and averages between 2,500 and 3,000 km [20]. Therefore, the driver interacts with the truck constantly and is the first person who is able to detect any potential technical deficiencies. The driver's response to unforeseen circumstances must therefore be timely and effective, which can significantly reduce potential costs. An unplanned stop at the roadside not only delays delivery but can also cause damage to the cargo [38].

In addition to requiring that the truck driver has the necessary professional competencies, transport companies should consider utilising the potential that drivers can represent in strategic fleet maintenance. Modern maintenance principles expect transport companies to integrate drivers into the entire maintenance system in order to take over parts of the professional competencies of maintenance workers. Transport companies must therefore ensure that drivers are adequately trained in the following aspects of maintenance: reactive action (the driver can remedy minor failure on the truck); preventive action (the driver inspects the truck before and after the trip is completed and notes any detected issues) and predictive-proactive action (the driver participates in the maintenance process and proposes corrective maintenance actions at company level).

But are heavy truck drivers prepared to do all this? Are international road transport companies capable of operating according to modern maintenance principles, much like manufacturing companies already do?

To answer these research questions, a study was conducted on a sample of 249 drivers in international road transport. The aim of the study is to determine how maintenance in transport companies affects the role of drivers in maintenance and the consequent optimal utilisation and roadworthiness of the trucks. To this end, the Structural Equation Model (SEM) was developed, which includes three research phases. The first phase consists in determining how maintenance in transport companies affects the reactive, preventive and predictive-proactive maintenance by truck drivers. The second phase consists in determining how individual driver maintenance affects truck utilisation. The third phase analyses the claim that the truck is in a good technical condition, despite the truck's high level of utilisation.

The results obtained build upon the current field of research into modern maintenance strategies, as they highlight the importance of establishing an efficient maintenance infrastructure in transport companies. This allows drivers to be actively engaged in the company's maintenance system, which ensures that the trucks are in a better technical condition, despite their high utilisation. The results therefore represent an important guideline for improving maintenance in transport companies, developing new competencies of truck drivers and upgrading existing knowledge in the development of modern operation maintenance. The article is structured as follows: the second part provides an overview of the literature in the field of maintenance, and an SEM model that was developed based on gaps in theory and practice. The research methods used are presented, followed by the results obtained on a sample of heavy truck drivers. The article continues with a discussion of the results, which highlights the contribution to maintenance theory and practice in transport companies. The conclusion highlights the limitations of the research and provides suggestions for further research.

2. Literature Review and Development of a Structural Research Model

The development of maintenance approaches allows for the distinction of three periods. The first period is characterised by reactive maintenance without prior planning. Repairs are only made when a failure occurs. According to Dekker [7], such an operational level of maintenance is difficult to manage due to possible unplanned events. If these events occur frequently, they interrupt important planned activities and prevent strategic thinking. This explains to some extent the general problem with maintenance dynamics, which typically overburdens the maintenance department with reactive work instead of proactive activities [16]. Although neglecting maintenance activities can have short-term effects, such as lower costs or more production hours, it can eventually cause problems resulting from more frequent failures, reduced capacity, or less time for continuous improvement [26].

That is why the second maintenance period is focused on preventive maintenance, with a focus on planned and preventive repairs. This enables companies to avoid unexpected component breakdowns based on a preventive maintenance plan [15]. However, Legutko [24] emphasises that the strategy of preventive operation maintenance is not successful when a significant number of failures occur at an early stage of utilisation. This fact and the increasing number of utilised machines/trucks in companies and devices of growing value have changed the attitude to this maintenance mindset.

The third period of maintenance therefore focused on predictive-proactive maintenance aimed at ensuring the smooth operation of machines and devices throughout their period of utilisation. This approach has given rise to various modern maintenance strategies, such as: TPM (total productive maintenance), RCM (reliability centred maintenance, RBI (risk-based inspection), CBM (condition-based monitoring), CMMS (computerised maintenance management software), RAM (Reliability, availability and maintainability program), OTI (Operator's own technical inspections), TSM (Technical systems maintenance), PDSS (Promise decision support system) and Lean principles.

The goal of predictive-proactive maintenance is to avoid as many system breakdowns as possible using historical data, empirical tests and statistical calculations. This puts into direct relation the time (working time and/or life span) with the probability of system breakdown [15]. However, the first condition for introducing modern maintenance strategies is to identify the maintenance function as an important strategic area within the company. Performing strategic operation maintenance within a company involves identifying and eliminating the basic losses that may occur: loss of availability (machine failures, adjustment and tool exchange), losses of performance (idling, reduced operating speed), and quality loss (rejects and modifications, losses due to test batches). Its implementation requires: the preparation of a program ensuring autonomous realisation of the operation maintenance works, planning of the activity for the organisational unit responsible for the operation maintenance works, improvement of the skills of the staff responsible for operation maintenance, and the preparation of a program oriented for optimisation of the work of new machines and devices [24].

The introduction of modern maintenance strategies also highlights a changed understanding of the organisation of maintenance, which leads to greater collaboration and teamwork in the company-wide maintenance process. An important feature of modern strategies is the introduction of independent machine and device maintenance by the operators, that is, connecting many basic service actions with the manufacturing process. This highlights the new role that operators are expected to play in relation to the machine that is not merely associated with production tasks but also with maintenance tasks. Initially, all observations of the functioning of machines and devices are performed by operators during their normal work, such as watching the instrumentation, day-to-day checking of predetermined elements with the use of physical effects such as smell, sounds, vibration, temperature, sight observation, changes of appearance, necessity to apply force etc. Then, after training, the operators perform inspections on their own, take full responsibility for the inspections, maintenance, cleaning, adjustments and small repairs [13, 24].

Therefore, the efficiency of the introduction of modern maintenance strategies largely depends on ensuring adequate resources for their implementation. The main factor that affects the quality of maintenance processes is competent employees. Their knowledge, skills and ability to respond to unexpected situations largely determine the efficiency of the functioning of the technical infrastructure in an enterprise [3]. Modern maintenance strategies therefore focus on two aspects of enhancing the competencies of employees: the first is to improve the competence of the staff responsible for operation maintenance. Another aspect is the engagement of operators in the works for operation maintenance and the transfer of responsibility to them allows for a better utilisation of the knowledge they possess, reinforces their sense of their own value and makes them aware of their participation in the achievement of the enterprise's objectives [24].

However, there is a fundamental difference between the maintenance of machinery in manufacturing and fleet maintenance in companies. While researchers and maintenance engineers in manufacturing companies are thoroughly invested in enhancing the advanced actions of operators in machine maintenance [3], the role of the driver in truck maintenance remains a neglected and under-researched field. Working in the transport industry has its own unique features, since the average driver in international transport spends five days a week in a truck and averages between 2,500 and 3,000 km [20]. Therefore, the driver is the first person who is in the position to detect any potential issues with the truck, even more so than the machine operator. The driver's response to unforeseen circumstances must therefore be timely and effective, which can significantly reduce potential costs. An unplanned stop at the roadside not only delays delivery but can also cause damage to the cargo [38].

That is why the truck driver must be properly trained. The requirements for truck drivers in terms of knowledge and skills in the area of maintenance do not differ in the EU [39] and in the USA [6]. In both cases, the service provided by the driver is defined as the transport and delivery of cargo to a specified location within the prescribed time. The driver is responsible for keeping the vehicle clean and regularly maintained. The driver must also have a vocational high school diploma. The US Department of Labor [6] adds that the drivers must inspect their trucks before and after the trip and record any defects they find. They must immediately report any technical problems to the competent staff.

However, the required knowledge and competencies of drivers may not suffice, if companies in the transport industry fail to utilise the potential that drivers can represent in the efficient maintenance of the fleet. As the driver is identified as the first person able to detect and correct potential defects on the truck, the aim of the study is to determine how maintenance in transport companies affects the role of drivers in truck maintenance.

To this end, the Structural Equation Model (SEM) was developed, which includes three research phases (Fig. 1). In the first phase, it is determined how maintenance in transport companies (MTC) affects truck drivers' ability to perform small repairs (SR), preventive maintenance (PRE) and proactive maintenance (PRO). The second phase continues by determining how individual maintenance by the driver affects truck utilisation (TU), measured in kilometres driven. The third phase concludes the study with an analysis of the claim that the truck is in a good technical condition (TCT), despite the truck's high level of utilisation. This is possible due to the coordinated work by all the persons involved in the company's maintenance process.

The SEM model is based on the following hypotheses:

H1: Maintenance in transport companies affects truck drivers' ability to perform small repairs.

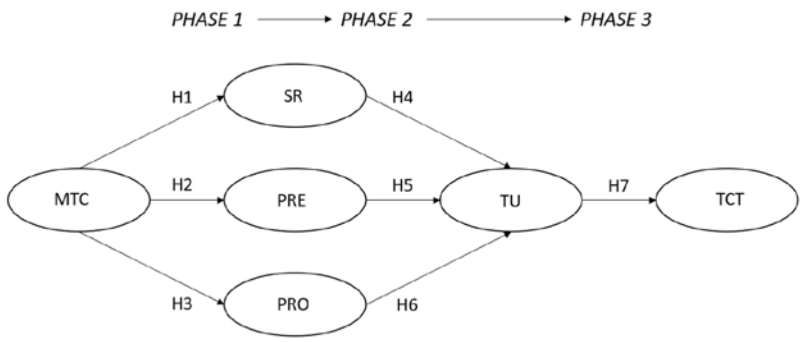


Fig. 1. Structural Equation Model

Much like machine operators, drivers in transport companies must also be trained to perform small repairs [1, 18], such as changing tires, oil, lights, and keeping the truck clean.

H2: Maintenance in transport companies affects the drivers' ability to take preventive action.

Much like machine operators, drivers in transport companies must also be trained to inspect the truck's equipment and supplies [16, 18], such as tires, lights, brakes, oil and water, in the context of preventative maintenance. In the event of mechanical problems, the driver must inform the appropriate personnel. The driver must prepare truck inspection reports and keep records of maintenance work/repairs performed.

H3: Maintenance in transport companies affects the drivers' ability to take proactive action.

Much like machine operators, drivers in transport companies must also be trained to plan maintenance, to participate in the maintenance process, and be encouraged to propose corrective actions at company level [25, 29], in the context of predictive-proactive maintenance.

H4: The ability of the driver to perform small repairs affects the truck's utilisation.

The driver immediately repairs the trucks when a minor failure occurs, instead of adding to the workload of the maintenance department. By doing so, drivers improve the availability and utilisation of the truck and avoid becoming standing by the road unable to continue the transport mission [38].

H5: The driver's ability to take preventive action affects the utilisation of the truck.

If the driver discovers technical issues with the truck before the start of the trip and after the trip, and promptly resolves them or informs the maintenance department in a timely manner, future failures on the road can be avoided. An unplanned stop not only delays delivery but can also cause damage to the cargo. Thus, an ability to detect early symptoms of wear, before they become real problems, has the potential to improve the truck's availability and utilisation [32].

H6: The drivers' ability to take proactive action affects the truck's utilisation

A high degree of availability can be achieved by frequently replacing components, but this approach can be expensive, not only because of frequent visits to the workshop, but also because of the cost of components. Therefore, failure prognostics and thinking about what needs to be improved in truck maintenance represent the highest level of development in maintenance by drivers [38]. Such driver action leads to greater engagement and teamwork in the maintenance process [28].

H7: Higher truck utilisation does not affect the truck's technical condition

The higher the level of utilisation of the vehicle, the greater the chance of failure and component wear. However, due to the effect of maintenance in transport companies on reactive, preventive and predictive-proactive maintenance by truck drivers, the truck remains in a good technical condition.

3. Methods

The study was conducted on a sample of 249 heavy truck drivers employed in the international road freight transport sector. For this purpose, an online questionnaire was prepared and sent via SMS to individual truck drivers at truck terminals in Koper (Slovenia) and Trieste (Italy). These two locations were selected due to the high number of trucks passing through and the proximity of the ports of Trieste (Italy) and Koper (Slovenia).

Due to the international nature of the study, the online questionnaire allowed the respondents to answer in Italian, Slovenian, Croatian, Serbian and English. Prior to the official submission of the questionnaire to the drivers, the questionnaire was examined by five heavy truck drivers who were not subsequently included in the study. Of the 731 drivers who were invited to the take part in the study, 249 drivers accepted the invitation and also completed the online questionnaire. 96 truck drivers were from Slovenia, 92 from Serbia, 37 from Croatia, 22 from Italy and 2 from Austria. The response rate was 34%. The final sample of respondents who completed the questionnaire in its entirety included 220 drivers. The study was divided into two parts. The first general part is connected to the demographic information such as the gender, age and education level of heavy truck drivers. The second part of the study focuses on thematic sections related to maintenance in companies, maintenance by drivers, technical characteristics and truck utilisation.

3.1. Measures

Maintenance in transport companies (Cronbach's α = 0.81) was measured by three indicators: "The company regularly conducts preemptive truck inspections every weekend", "The company ensures that every technical issue of the vehicle is quickly fixed" and "The company provides me with training about basic vehicle maintenance procedures".

Preventive behaviour of the truck drivers (Cronbach's α = 0.82) was measured by two items: "Before I start driving, I always perform a "walk around" check to spot potentially dangerous issues" and "I immediately notify the company in case of any issues spotted during the maintenance of the truck" Cronbach's α equals 0.82.

Proactive behaviour of the truck drivers was a single item measure "I often think about what needs to be improved on truck maintenance".

Inclination to do small truck repairs (Cronbach's α = 0.85) was measured by two items: "If necessary, I change the headlight bulb of the truck" and "I regularly clean the cab".

Technical condition of the truck (Cronbach's α = 0.78) was measured by two items: "The truck I'm driving is in good condition, based on visual inspection" and "The truck is comfortable to drive".

Truck utilisation was a single-item measure. Respondents answered the question "What is your average number of kilometres driven per week?".

3.2. Procedure

Data analysis followed the two-step approach suggested by Anderson and Gerbing [2], evaluating measurement model via exploratory and confirmatory factor analysis in the first step and building a structural equation model in the second. Robust maximum likelihood method of parameter estimation as proposed by Boomsma & Hoogland [5] for data not following multivariate normal distribution was used for the evaluation of measurement and for the structural model. The error variance of the constructs with a single indicator was set to 0.

If each observable variable loaded statistically significantly onto the factor it was supposed to measure, this was considered a sign of convergent validity [2, 19, 34, 37]. Furthermore, item loadings on the factor were supposed to be higher than 0.5 [21, 34]. The sign of convergent validity was also a good overall fit of the model [34].

Composite reliability and average variance extracted (AVE) were calculated. The first is the reliability of a summated scale and the second the variance in the indicators explained by the common factor. The composite reliability above 0.6 indicates good reliability and AVE above 0.5 good construct validity [14]. Reliability was assessed also by Cronbach's α. The values above 0.70 were considered to indicate adequate reliability as proposed by Nunnally [31].

Discriminant validity was assessed by examining correlations between constructs. As a rule of thumb, correlation between latent variables should not exceed r = 0.85, otherwise one can conclude that there is a lack of discriminant validity [23]. Discriminant validity is supported further if the 95% confidence interval for the correlation coefficient between two latent variables does not include 1 [35]. If the AVE by the correlated latent variables is greater than the square of the correlation between the latent variables, then discriminant validity exists [14].

After establishing good measurement validity, a structural equation model was built. Partial disaggregation approach was used due to the rather small sample size in which subsets of items were combined into composites (averages) that were then treated as indicators of the constructs. Error variances of the indicators were calculated as (1-reliability)*variance of the indicator.

The fit of the model was evaluated using the Sattora-Bentler scaled Chi Square, which is suitable for evaluating models with non-normal data [5]. In addition, the Comparative Fit Index (CFI), Incremental fit index (IFI), non-normed fit index (NNFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardised Root Mean Square Residual (SRMR) were used. Values of .95 or and above or 0.90 and above for CFI, NNFI and IFI, and values of 0.08 and below for RMSEA and SRMR indicate a good fit of the model [22]. Modification indices and standardised residuals were used to suggest improvements of the model.

LISREL 9.30 was used for model calibration and hypotheses testing.

4. Results

The final sample included n = 220 truck drivers. One woman participated in the study, while other participants were males. The average age (SD) of participants was 40.8 (10.7) years. Approximately one third (37.7 %) had more than high-school education. The average (SD) number of years working as a truck driver was 13.7 (10.1) years. On average (SD), they are employed at the current employer for 5 (5.1) years.

Confirmatory factor analysis showed good overall fit of the measurement model (SB χ^2 = 35.02; df = 31; p = 0.283; SB χ^2/ df = 1.1; RMSEA = 0.059; NFI = 0.97; NNFI = 0.99; CFI = 0.99; IFI = 0.99; SRMR = 0.04; GFI = 0.96). Convergent validity was further supported by significant and substantial (> 0.50) loadings for all items in the model. AVE for all constructs was 0.60 or higher for all constructs, supporting construct validity.

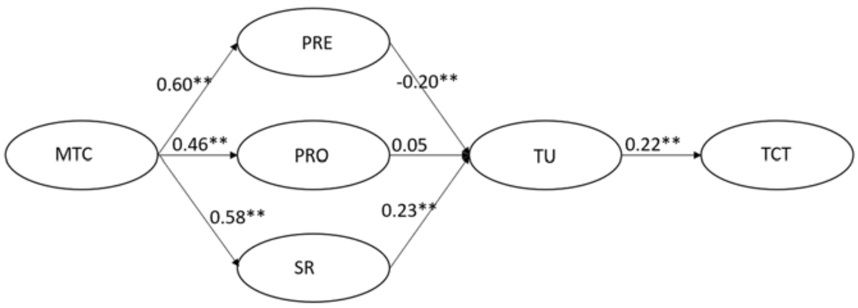
Descriptive statistics for composite variables (averages of the indicators measuring each construct) and correlations are presented in Table 1. None of the correlation coefficients exceeds 0.85 and none of their confidence intervals include 1. None of the squared correlation coefficients exceeds AVE. From the stated, it is concluded that there is good discriminant validity of the used measures.

The hypothesis testing included building proposed structural equation model. The structural equation model with standardised regression coefficients is depicted in Fig. 2. All presupposed paths are statistically significant but the one from proactive behaviour to truck

Table 1. Correlation matrix of composite measures

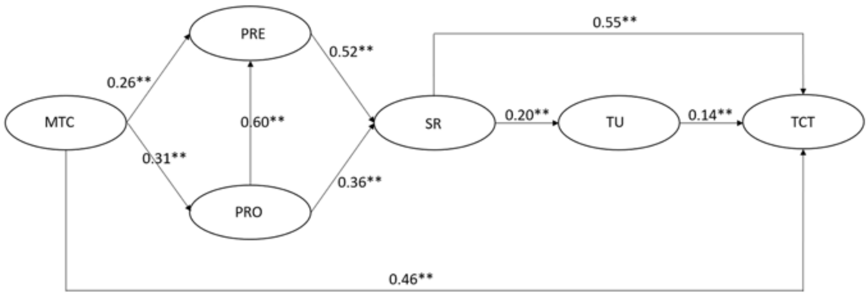
	M	SD	PRE	SR	MTC	TCT	PRO	TU
PRE	4.44	0.71	1					
SR	4.31	0.85	0.65**	1				
MTC	3.87	0.95	0.37**	0.36**	1			
TCT	4.41	0.74	0.45**	0.63**	0.53**	1		
PRO	4	1.01	0.62**	0.65**	0.28**	0.51**	1	
TU	2561.5	690.4	0.05	0.20**	−0.05	0.20**	0.12	1

** p < 0.05; ** p < 0.01; MTC = maintenance in transport companies; PRE = preventive behaviour; PRO = proactive behaviour; SR = small repairs; TCT = technical condition of the truck; TU = truck utilisation*



**p < 0.05; ** p < 0.01*

Fig. 2. Structural equation model (standardised regression coefficients are shown; MTC = maintenance in transport companies; PRE = preventive behaviour; PRO = proactive behaviour; SR = small repairs; TCT = technical condition of the truck; TU = truck utilisation)



**p < 0.05; ** p < 0.01*

Fig. 3. Modified structural equation model (standardised regression coefficients are shown; MTC = maintenance in transport companies; PRE = preventive behaviour; PRO = proactive behaviour; SR = small repairs; TCT = technical condition of the truck; TU = truck utilisation)

utilisation. The fit of the model, however, is very poor (SB $\chi^2 = 113.5$; df = 8; $p < 0.001$; SB χ^2/ df = 14.2; RMSEA = 0.45; NFI = 0.77; NNFI = 0.58; CFI = 0.78; IFI = 0.78; SRMR = 0.27; GFI = 0.63).

Due to poor fit of the presupposed structural model, the modification of the model was done and the modified model is depicted in Fig. 3. The model exhibits good overall fit (SB $\chi^2 = 8.6$; df = 6; $p < 0.200$; SB $\chi 2$/ df = 1.4; RMSEA = 0.083; NFI = 0.98; NNFI = 0.99; CFI = 0.99; IFI = 0.99; SRMR = 0.04; GFI = 0.98). It presupposes full mediation of the proactive and preventive behaviour of the driver between the maintenance in transport companies and small repairs done by the driver. Partial mediation through truck utilisation is present between small repairs done by the driver and the technical condition of the truck. Maintenance in transport companies influences the technical condition of the truck directly, as well as through the driver's preventive and proactive behaviour and truck utilisation.

5. Discussion

The results of the study indicate that truck drivers are able to participate in the maintenance process in transport companies, as they are trained to perform minor maintenance operations in the same way as machine operators [1, 18]. Drivers can also be steered toward preventive thinking, as they immediately report any defects on the truck and always perform a 'walk around' check to spot potential issues. The same goes for proactive action, as drivers often think about what needs to be improved in truck maintenance. This is also in line with current operational maintenance guidelines, where part of the maintenance personnel tasks is transferred to the direct users of machines or devices [24]. Based on this, the first research question can be answered, since the basic results of the study indicate that truck drivers are willing and able to actively participate in the maintenance system of transport companies. They are able to perform minor maintenance work and act both preventively and proactively. The results therefore build upon the set of professional competencies as indicated in the recommendations by the Ministries of labour and employment in the EU [39] and in the US [6] and highlight the new role that truck drivers must play in operation maintenance. This role consists in an active involvement in the whole process of fleet maintenance in transport companies. The results also add to the official maintenance standards in the transport industry, developed by the British DVSA - Driver and Vehicle Standards Agency [12], where the tasks of the driver are limited to daily routine truck inspection and prompt reporting and recording of all technical defects [27].

However, an important aspect of building an effective maintenance infrastructure consists primarily in determining whether transport companies are capable of operating according to modern operating maintenance principles, as is common in the manufacturing industry. This aspect was addressed by the second research question. The basic results of the survey, i.e. the median value, showed that, on average, the companies in the sample have a well-organised system of fleet maintenance. However, in order to determine how maintenance in transport companies affects the role of the drivers in the maintenance and consequent utilisation and good technical condition of trucks, an SEM model was developed and divided into the three research phases, as shown in Fig. 1. The results of the first research phase showed that the effective functioning of operation maintenance in transport companies affects the ability of truck drivers to operate within the framework of preventive and proactive maintenance. The results are also in line with modern preventive maintenance guidelines in the manufacturing industry, according to which machine operators can only act preventively and proactively, if the companies provide an adequate maintenance infrastructure: from enhancing the maintenance competencies of users of machinery or transport vehicles [3] to the introduction of modern operation maintenance processes [25].

Transport companies that are systematically focused on fleet maintenance enable drivers to participate in the maintenance process and encourage them to propose corrective actions. As indicated by the results of the final part of the first research phase, truck drivers who are proactive also tend to act preventively. This allows them to be able to repair minor failures, resulting in higher truck utilisation and more kilometres driven. This is also the final result of the second research phase, in which the goal was to determine how individual maintenance by the driver affects truck utilisation, as measured in kilometres driven. The results of the second research phase point out that, by doing so, drivers improve the availability and utilisation of the truck and avoid becoming standing by the road unable to continue the transport mission [38]. In addition, the ability to detect early symptoms of wear, before they become real problems, has the potential to improve the truck's availability and utilisation, and the drivers' ability to take timely action [32].

The third research phase of the SEM model rounded off the entire research field of operation maintenance, both in comparison with the manufacturing activity, as well as by setting future guidelines for operation maintenance in the transport industry. The third phase concludes the study with an analysis of the claim that the truck is in a good technical condition, despite the truck's high level of utilisation. The results showed that this is only possible with the coordinated action of all the persons involved in the company's maintenance process. The first condition is that companies must have an effective operation maintenance infrastructure in place, which enables them to conduct a preventative inspection of the truck every weekend, to quickly eliminate any technical defects on the vehicle and to provide truck drivers with training on the basic procedures of vehicle maintenance. As a consequence, the drivers will be more proactive, which will affect their preventive action. This will enable them to repair minor failures, which will result in better utilisation of the truck. Despite being utilised more, the trucks will remain in a good technical and visual condition, which is the most important finding of the research. The final finding therefore highlights the importance of establishing an adequate maintenance infrastructure in the company and, above all, how it affects the participants in the maintenance process.

5.1. Contribution to Theory and Practice

The results can be generalised to the entire transport industry, as the study was conducted on a sufficiently large number of truck drivers. Several studies have focused on the work of truck drivers in the transport industry, but all were limited to researching the labour law aspect (regulations regarding driving time, rest time, hours of service violations, etc.), driver health issues (fatigue, general health, etc.), and driver behaviour in traffic (research via DBQ - Driver Behaviour Questionnaire). The study thus represents an original scientific novelty, as it is the first study that examines the position of maintenance in the transport industry and in particular how maintenance in transport companies affects the role of drivers in maintenance and the consequent optimal utilisation and roadworthiness of the trucks. The research gap is not only noticeable in scientific publications but also in EU laws, since Directive 2003/59/EC [11], which defines the procedures for obtaining a basic professional qualification, Directive 2014/45/EU (on periodic roadworthiness tests for motor vehicles and their trailers) [8], Directive 2014/46/EU (on the registration documents for vehicles) [9], and Directive 2014/47/EU (on the technical roadside inspection of roadworthiness of commercial vehicles) [10], do not specifically define the role of drivers in fleet maintenance. Unlike in the transport industry, the field of maintenance in the manufacturing industry is undergoing continuous development in research, since the first studies on modern approaches to maintenance date back to the 1970s [24].

The innovativeness of the results therefore affects the following theoretical, practical and legislative areas of maintenance in the transport industry:

- Truck maintenance: It was determined that drivers play an important role in fleet maintenance. Drivers are capable of repairing minor failures, but only if they are proactive, which will affect their preventive behaviour. This can greatly benefit transport companies, as the results show that the vehicles are utilised more efficiently and in a better technical condition. The main novelty consists in the fact that truck utilisation is affected not only by the economical driving of truck drivers, as pointed out by Lowe & Pidgeon [27] and Muha & Sever [30], but also by their active involvement in the company's maintenance process.
- Development of truck drivers' competencies: The requirements for truck drivers regarding professional training in the EU are limited to the basic qualification - code 95. New legislation in the road freight transport sector requires that all new drivers must obtain a basic qualification - code 95, which they obtain on the basis of a theoretical and practical exam before a state commission [11, 34]. The results highlight the importance of training truck drivers in the field of vehicle fleet maintenance, as transport companies that operate in this way also have vehicles that are in better technical condition. Knowledge of fleet maintenance is not specifically covered by code 95 training, thus the study highlights further directions in the training of truck drivers and expands the required competencies both in the practical and in the legislative field.
- Maintenance improvements in transport companies: It was determined that companies will only be able to take advantage of the driver's maintenance potential, if they have an effective operation maintenance infrastructure in place, which enables them to conduct a preventative inspection of the truck every weekend, to quickly eliminate any technical defects on the vehicle and to provide truck drivers with training on the basic procedures of vehicle maintenance. Since this area is unresearched, the results represent an important guideline for establishing the strategic role of maintenance in transport companies and the efficient organisation of maintenance at all levels in the company. This is especially important since the field of operation maintenance development in the transport industry is lagging behind the manufacturing sector, where the guidelines for the operation of the maintenance function are clearly defined within Industry 4.0 [33].
- The results build upon existing knowledge in the field of modern operation maintenance development, as they show how the maintenance function is tied into a series of interconnected maintenance actions in transport companies. The responsibility for the development of maintenance infrastructure and efficient operation maintenance, which ultimately leads to a better condition of vehicles, therefore depends on the management of the company responsible for establishing an adequate maintenance infrastructure. The results are therefore consistent with the study by Bokrantz et al. [4], where they highlight scenarios for future modern maintenance management by 2030. The study emphasises that corporate leadership should be aware of the fact that failure to develop an adequate level of competence is linked to an inefficient maintenance process, which in turn increases the sensitivity to disturbances, decreases the responsiveness to failures, and at the same time reduces the competitiveness of the company. Therefore, education and training of employees is an absolute necessity, as well as developing new ways of managing competencies in the company maintenance process. While Bokrantz et al. [4] highlighted guidelines for future operation maintenance by 2030, some of their recommendations have been explicitly confirmed by the results of the presented study.
- The study synthesises and analyses the literature in the field of maintenance development through different maintenance periods, which is an important framework for future empirical studies in the transport industry, as well as in the field of operation maintenance in general.

Although the survey included truck drivers from Slovenia, Italy, Serbia and Croatia, the results are representative for the whole EU area and not just for the Central European region or for each individual country. This assertion can be justified by the fact that all truck drivers driving in the EU operate within the framework of the uniform law provisions related to working time, obligatory rest periods for mobile workers and recording equipment used in road transport [30]. Furthermore, the transport companies where truck drivers work must ensure that their fleet is in compliance with uniform EU legislation [8-10]. As truck drivers operate according to uniform rules, it is possible to draw global-scale conclusions from the study and highlight suggestions for improvements in operational maintenance in the transport industry, which are the following:

- Improving the level of operational maintenance in transport companies,
- Training of drivers for basic qualification (code 95), which is a prerequisite for drivers to operate within the EU, should include training in fleet technical maintenance,
- EU legislation should state that truck drivers must also be responsible for the roadworthiness of trucks.

5.2. Limitations and Recommendations for Future Studies

This is the first study that deals with the field of operation maintenance in the transport sector. Therefore, it was not possible to make a substantial comparison of the results with previous studies. In this regard, the limitations of the study also dictate recommendations for future studies, since it would be advisable to repeat a similar study in the future and compare the evolution of the maintenance function in the transport industry. As regulations governing the work of drivers differ around the world, which can affect their ability to maintain the fleet, it would be advisable to conduct a related study in the US, in Australia, in Canada and in other countries.

In the future, the field of operation maintenance should be regulated in more detail within the EU legislation (via system instructions and procedures), since there are no specific guidelines as to who is in charge of managing maintenance procedures in a transport company and how these procedures should be managed. EU legislation [8-10] states that keeping the truck in a roadworthy condition is solely the responsibility of the transport company and this is as far as this area of maintenance is addressed. Systematic truck maintenance is especially important from a safety point of view, since a road accident that results in a technically defective truck often has more serious consequences than if caused by other vehicles. Therefore, it is essential that other subjects in transport companies are included in the legislation that regulates responsibilities in operation maintenance. The study results show that drivers can significantly contribute to the roadworthiness of the fleet. Unlike continental Europe, in the UK, the DVSA [12] defines, through clear procedures, the driver's responsibility for operation maintenance. In these procedures, drivers are required to formally record all defects on the truck, any repairs made and the signature of the person who carried out the repair. Drivers are also required to indicate how much time they spend on daily 'walk around' checks. The drivers' tachographs must clearly show the time spent on the checks performed before driving on the road. This is not clearly defined and regulated in the continental EU.

The following recommendation for further research is related to the future development of trucks and driver-assistance systems that change the way driving is done, as well as truck maintenance procedures. The competencies required from drivers will change accordingly, and both research and modern recommendations for operation maintenance from a technical, organisational and human resources development point of view will have to follow.

6. Conclusions

While there have been many studies on the subject of the transport of goods in international transport, this is the first study to examine how maintenance in transport companies affects the role of drivers in fleet maintenance. The results obtained build upon the current field of research into modern maintenance strategies, as they highlight the importance of establishing an efficient maintenance infrastructure in transport companies. This allows drivers to be actively engaged in the company's maintenance system, which ensures that the trucks are in a better technical condition, despite their high utilisation. The results therefore represent an important guideline for improving maintenance in transport companies, developing new competencies of truck drivers and upgrading existing knowledge in the development of modern operation maintenance.

Although this is the first study conducted in the field of operation maintenance in the transport industry, the research has its limitations, tied to the inability to compare the results with previous studies. This limitation also dictates future recommendations, since it would be advisable to repeat a similar study in the future and compare the evolution of the maintenance function in the transport industry.

References

1. Andersson R, Manfredsson P, Lantz B. Total productive maintenance in support processes: an enabler for operation excellence. Total Quality Management & Business Excellence 2015; 26(9-10): 1042-1055, https://doi.org/10.1080/14783363.2015.1068598.
2. Anderson J C, Gerbing D W. Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. Psychological Bulletin 1988; 103(3): 411-23, https://doi.org/10.1037/0033-2909.103.3.411.
3. Antosz K. Maintenance - identification and analysis of the competency gap. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2018; 20(3): 484-494, https://doi.org/10.17531/ein.2018.3.19.
4. Bokrantz J, Skoogh A, Berlin C, Stahre J. Maintenance in digitalised manufacturing: Delphi-based scenarios for 2030. International Journal of Production Economics 2017; 191: 154-169, https://doi.org/10.1016/j.ijpe.2017.06.010.
5. Boomsma A, Hoogland J J. The robustness of LISREL modeling revisited. In R. 2001.
6. Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook, Heavy and Tractor-trailer Truck Drivers. Available at https://www.bls.gov/ooh/transportation-and-material-moving/heavy-and-tractor-trailer-truck-drivers.htm (visited July 19, 2019).
7. Dekker R. Applications of maintenance optimization models: a review and analysis. Reliability Engineering & System Safety 1996; 51(3): 229-240 https://doi.org/10.1016/0951-8320(95)00076-3.
8. Directive 2014/45/EU of the European Parliament and of the Council of 3 April 2014 on periodic roadworthiness tests for motor vehicles and their trailers and repealing Directive 2009/40/EC Text with EEA relevance. Available at http://data.europa.eu/eli/dir/2014/45/oj, (visited December 19, 2019).
9. Directive 2014/46/EU of the European Parliament and of the Council of 3 April 2014 amending Council Directive 1999/37/EC on the registration documents for vehicles. Available at http://data.europa.eu/eli/dir/2014/46/oj, (visited December 19, 2019).
10. Directive 2014/47/EU of the European Parliament and of the Council of 3 April 2014 on the technical roadside inspection of the roadworthiness of commercial vehicles circulating in the Union and repealing Directive 2000/30/EC. Available at http://data.europa.eu/eli/dir/2014/47/oj, (visited December 19, 2019).
11. Directive 2003/59/EC of the European Parliament and of the Council of 15 July 2003 on the initial qualification and periodic training of drivers of certain road vehicles for the carriage of goods or passengers, amending Council Regulation (EEC) No 3820/85 and Council Directive 91/439/EEC and repealing Council Directive 76/914/EEC. Available at http://data.europa.eu/eli/dir/2003/59/2019-07-26. (visited December 19, 2019).

12. Driver and Vehicle Standards Agency, DVSA, 2019. Available at https://www.gov.uk/government/organisations/driver-and-vehicle-standards-agency. (visited September 25, 2019).
13. Epler I, Sokolović V, Milenkov M, Bukvić M. Application of lean tools for improved effectiveness in maintenance of technical systems for special purposes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2017; 19(4): 615-623, https://doi.org/10.17531/ein.2017.4.16.
14. Fornell C, Larcker D F. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research 1981, 18(1): 39-50, https://doi.org/10.2307/3151312.
15. Fornasiero R, Zangiacomi A, Sorlini M. A cost evaluation approach for trucks maintenance planning. Production Planning & Control 2012; 23(2-3): 171-182, https://doi.org/10.1080/09537287.2011.591641.
16. Gania I P, Fertsch M K, Jayathilaka K R K. Reliability Centered Maintenance Framework for Manufacturing and Service Company: Functional Oriented. 24th International Conference on Production Research 2017: 721-725.
17. Groebler A, Thun J H, Milling P M. System dynamics as a structural theory in operations management. Production and Opererations Management 2008; 17(3): 373-384, https://doi.org/10.3401/poms.1080.0023.
18. Habidin N, Hashim S, Fuzi N, Salleh M. Total productive maintenance, kaizen event, and performance. International Journal of Quality & Reliability Management 2018; 35(9): 1853-1867, https://doi.org/10.1108/IJQRM-11-2017-0234.
19. Hair J, Anderson R, Tatham R, Black W. Multivariate Data Analysis (5th ed.). London, Prentice Hall International. 1998.
20. Hege A, Perko M, Johnson A, Yu C H, Sönmez S, Apostolopoulos Y. Surveying the Impact of Work Hours and Schedules on Commercial Motor Vehicle Driver Sleep. Safety and Health at Work 2015, 6(2): 104-113, https://doi.org/10.1016/j.shaw.2015.02.001.
21. Hildebrandt L. Consumer retail satisfaction in rural areas: A reanalysis of survey Data. Journal of Economic Psychology 1987; 8(1): 19-42, https://doi.org/10.1016/0167-4870(87)90004-3.
22. Hu L, Bentler P M. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods 1998; 3(4): 424-453, https://doi.org/10.1037/1082-989X.3.4.424.
23. Kline R B. Principles and practice of structural equation modeling. New York: Guilford Press. 2011
24. Legutko S. Development trends in machines operation maintenance. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2009; 42(2): 8-16.
25. Levitt J, Complete Guide to Preventive and Predictive Maintenance. Industrial Press Inc, New York. 2011.
26. Linneusson G, Ng H C, Amos, A T. Towards strategic development of maintenance and its effects on production performance by using system dynamics in the automotive industry. International Journal of Production Economics 2018; 200: 151-169, https://doi.org/10.1016/j.ijpe.2018.03.024.
27. Lowe D. Pidgeon, C. Lowe's Transport Manager's and Operator's Handbook. Kogan Page; 49th edition, 2019.
28. Moniri Morad A, Pourgol-Mohammad M, Sattarvand J. Application of reliability-centered maintenance for productivity improvement of open pit mining equipment: Case study of Sungun Copper Mine. Journal of Central South University of Technology 2014; 21(6): 2372-2382, https://doi.org/10.1007/s11771-014-2190-2.
29. Moubray J. Maintenance management - a new paradigm. Aladon Ltd, 1996.
30. Muha R, Sever D. The impact of regulation 561/2006 on fleet management viewed through efficient use of drivers' working time. Promet-Traffic&Transportation 2009; 21(1): 61-67, https://doi.org/10.7307/ptt.v21i1.913.
31. Nunnally J. Psychometric Theory (2nd ed.). New York: McGraw-Hill.of Research in Marketing, 1978.
32. Rögnvaldsson T, Nowaczyk S, Byttner S, Prytz R. Svensso M. Self-monitoring for maintenance of vehicle fleets. Data Mining and Knowledge Discovery 2018; 32(2): 344-384, https://doi.org/10.1007/s10618-017-0538-6.
33. Roy R , Stark R, Tracht K, Takata S, Mori M. Continuous maintenance and the future - Foundations and technological challenges. CIRP Annals -Manufacturing Technology 2016; 65: 667-688, https://doi.org/10.1016/j.cirp.2016.06.006.
34. Steenkamp J B, van Trijp H. Attribute Elicitation in Marketing Research: A Comparison of Three Procedures. Marketing Letters 1997; 8(2): 153-165, https://doi.org/10.1023/A:1007975518638.
35. Torkzadeh G, Koufteros X, Pflughoeft K. Confirmatory analysis of computer self-efficacy. Structural Equation Modeling: A Multidisciplinary Journal 2003; 10(2): 263-275, https://doi.org/10.1207/S15328007SEM1002_6.
36. Uradni list RS, št. 103/10, 63/12, 62/14 in 69/19. Pravilnik o temeljnih kvalifikacijah za voznike motornih vozil v cestnem prometu. Available at http://www.pisrs.si/Pis.web/pregledPredpisa?id=PRAV10254. (visited July 31, 2019).
37. Viera A L. & SpringerLink (Online Service). Interactive LISREL in Practice: Getting started with a SIMPLIS Approach. Heidelberg ; New York: Springer, 2011, https://doi.org/10.1007/978-3-642-18044-6.
38. Voronov S, Frisk E, Krysander M. Data-Driven Battery Lifetime Prediction and Confidence Estimation for Heavy-Duty Trucks. IEEE Transactions on Reliability 2018; 67(2): 1-17, https://doi.org/10.1109/TR.2018.2803798.
39. Zavod RS za zaposlovanje, 2019. Available at https://www.ess.gov.si/ncips/cips/opisi_poklicev/opis_poklica?Kljuc=2215&Filter. (visited July 19, 2019).

Sebastjan ŠKERLIČ
University of Ljubljana
Faculty of Maritime Studies and Transport
Pot pomorščakov 4, 6320 Portorož, Slovenia

Vanja ERČULJ
University of Maribor
Faculty of Criminal Justice and Security
Kotnikova ulica 8, 1000 Ljubljana, Slovenia

Edgar SOKOLOVSKIJ
Vilnius Gediminas Technical University
Faculty of Transport Engineering
J. Basanavičiaus 28, LT-03224 Vilnius, Lithuania

E-mails: sebastjan.skerlic@fpp.uni-lj.si, edgar.sokolovskij@vgtu.lt, vanja.erculj@fvv.uni-mb.si

Article citation info:
PASZKOWSKI W. Modeling of vibroacoustic phenomena using the method of parameterizing the audio signal. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 501–507, http://dx.doi.org/10.17531/ein.2020.3.13.

Waldemar PASZKOWSKI

MODELING OF VIBROACOUSTIC PHENOMENA USING THE METHOD OF PARAMETERIZING THE AUDIO SIGNAL

MODELOWANIE ZJAWISK WIBROAKUSTYCZNYCH Z ZASTOSOWANIEM METODY PARAMETRYZACJI SYGNAŁU FONICZNEGO*

The article proposes an original way of modeling vibroacoustic phenomena of exploited machines/devices using the method of audio parameterization. This method extends the current approach to this type of research and consists in taking into account the psychoacoustic effects associated with the emission of vibroacoustic energy. The proposed solution is based on the determination of mel-cepstral coefficients of the examined signal and its classification, due to the impact of noise. It was presented verification of the method on the example of studies on the impact of road noise sources.

Keywords: *vibroacoustic phenomena, signal parameterization, noise perception, modeling.*

W artykule zaproponowano oryginalny sposób modelowania zjawisk wibroakustycznych eksploatowanych maszyn/urządzeń z zastosowaniem metody parametryzacji sygnału fonicznego. Sposób ten rozszerza dotychczasowe podejście do tego rodzaju badań i polega na uwzględnianiu efektów psychoakustycznych towarzyszących emisji energii wibroakustycznej. Proponowane rozwiązanie opiera się na wyznaczeniu współczynników mel-cepstralnych badanego sygnału i jego klasyfikacji, ze względu na oddziaływanie hałasu. Przedstawiono weryfikację zastosowania metody na przykładzie badań oddziaływania źródeł hałasu drogowego.

Słowa kluczowe: *zjawiska wibroakustyczne, parametryzacja sygnału, percepcja hałasu, modelowanie.*

1. Introduction

The source of phenomena occurring during the exploitation of machines and devices are complex vibroacoustic processes. These phenomena consist of varied impacts of vibration, noise, air and material sounds, or medium pulsation in the areas of machine parts. These impacts are emitted to the environment in the form of vibroacoustic energy. The effects of vibroacoustic phenomena are the response of machines/devices in relation to their parts and environment. Carrier of information about vibroacoustic phenomena there is a signal, which may be subject to various transformations. Analysis of the vibroacoustic signal as a way of processing data into useful information is used not only in maintenance tasks [3, 10], diagnosis [13], reliability [9], assessment of the exploitation condition of technical objects [14], but also in the tasks of the impact of machinery/devices on the environment. This impact can be analyzed in open spaces, rooms and in relation to the human body. Vibroacoustic phenomena are all vibration and acoustic waveforms, which are related in a causal way. It was found, that in the study of these phenomena, the following issues should be included [5]:

- time and spatial distribution of characteristics, describing the energy coming from the source,
- system response in the form of a vibroacoustic transition function,
- interdependence between sources.

Vibroacoustic phenomena are described with using basic physical quantities such as sound pressure, speed, acceleration, displacement, force [5]. In turn, the vibroacoustic signal can be represented by a unary function or a vector [1].

The applied methods of analyzing vibroacoustic signals use various Fourier transforms, so you can get amplitude, phase or energy spectra. Fourier transform allows obtaining valuable information by changing the signal from the time domain to the frequency domain, especially when there is a high dynamics of change of signal parameters time. Vibroacoustic research require not only measurements of vibration and noise characteristics, but also obtaining information about the studied phenomenon. The assumed research goal requires specifying the measurement methodology in relation to the technical object or its environment [5].

2. Modeling of vibroacoustic phenomena and sound perception review of current solutions

Modeling of vibroacoustic signal are subject to vibration and acoustic phenomena, generated during the exploitation of machines/devices. For analyzing vibroacoustic processes in the environment, it is used acoustic modeling of machines and devices [5]. For this purpose, methods of modeling the sound field of machines/devices are used to identify sound sources. Characteristics of sound sources can be assessed by [5]:

- a sound field generated by the source,
- a source as an emitter of vibroacoustic energy.

Methods of modeling of vibroacoustic phenomenon are widely used in tasks including identification and reduction of vibration and noise sources, analysis of vibroacoustic processes [11], the assessment of the technical condition of machines/devices and their reliability, damage analysis of machine/device elements or the assessment of the impact of vibroacoustic energy on the human body. For example, diagnostic modeling methods use a vibroacoustic signal to examine for wear or damage to technical object parts [5]. Among the known methods of modeling these phenomena, the following methods can be distinguished: pressure, intensity, reciprocity, finite elements, inversion, Fourier transforms, short-time Fourier transformation, wavelet transformation [1, 5].

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

In the analysis of vibroacoustic phenomena, most often there are processed data and information directly related to the sound source, human or environment, including interactions between them. One of the goals of searching for the relationship between the characteristics of a machine/device in the environment (noise source) and the parameters of the acoustic field is the formulation of sound emission models for the purposes of assessing its impact [23]. In this regard, based on the vibroacoustic signal, appropriate physical characteristics of measured energy quantities are determined. On this basis it is defined the degree of harmful vibration or noise, negative impact of vibroacoustic phenomena on the human body, or on machine/device components or the environment. Input data for modeling vibroacoustic phenomena are acquired acoustic information from specified points of the machinery/device environment. For the purpose of determining the degree of exposure at a given point, it should be made measurements of characteristics of the noise pollution [6]. As a result of acoustic measurements, appropriate energy indicators are obtained at a given point [4, 21]. Studies on the auditory effects of human noise confirm, that noise risk assessment using only energy indicators is limited and insufficient [24, 20]. In the used approach, the subjective significance of noise impressions is ignored. It is not included interaction of sound pitch and the loudness of dynamically changing sounds. The A-, B, C- and D- filters, used in this respect, approximate the inverted shape of curves with equal loudness at different levels of sound pressure. Weighing with the A filter has become the most commonly used frequency factor, although it is not optimal for all sound pressure levels. Studies on sound perception show that the harmonics of sound with frequencies in the range 1÷5 [kHz] are more audible than the others [16]. This is important in interpreting the occurrence of elementary and complex phenomena in low and high frequency ranges [15]. In the research on the assessment of the subjective perception of human loudness for continuous sounds, it is used the method consisting in determining the total (equivalent) sound level corrected by a hearing correction filter (A). It should be noted that the frequency response of the filter (A) is an approximation of receiving audio impression of low sound levels. The analysis of the characteristics of the hearing threshold shows that hearing sensitivity is highest in the medium frequency range 1000÷4000[Hz] and it significantly decreases in the low and high frequency band. It should be noted that the characteristics of the thresholds of hearing, discomfort and pain significantly differ from each other as a function of frequency [12]. Analyzing the mechanism of acoustic impressions in relation to the characteristics of sound, it should be said, that there is no simple and clear relationship between the physical and subjective characteristics of sound. Research on subjective sound characteristics relate to the description and interpretation of auditory impressions. The results confirm that noise perception is significantly affected by among others psychoacoustic aspects of sound, signal time structure, shaping of subjective sound features in the domain of: time, frequency, time-frequency [8, 17, 18, 19]. The undertaken research on the assessment of the quality of audio signals is based on the use of two categories of methods [26]:

- subjective methods - consisting in assessing the human auditory impressions,
- objective methods - consisting in using approximate mathematical models to include perception mechanism.

Subjective methods of assessing sound quality use auditions of the assessed signal for the purpose of determining, perceived by the listener, the degree of similarity of both signals, the degree of their difference, or the level of discomfort caused by the presence of interference or distortion. These methods allow the evaluation of a direct response from the recipient, and individual subjective factors significantly affect a single assessment. Among the objective methods it can be distinguished:

- signal methods - most often the assessed signal is compared with the original signal, without distortion (reference signal). It is also used methods which do not take into account the reference signal.
- parametric methods - the quality of sound is assessed on the basis of knowledge about the applied processing technique and knowledge of its parameters, which are the input arguments of the assessment algorithm.

The review and interpretation of models and methods for analyzing vibroacoustic phenomena and sound quality assessment shows that the features contained in the audio signal can be a valuable source of information in the study of noise perception. It is important to take into account the psychoacoustic effects of the impact of vibroacoustic energy on the human body, both during its implementation of operational activities, maintenance, as well as being near sources of vibration and noise emissions.

The problem undertaken by the author concerns the inclusion, in modeling tasks of vibroacoustic phenomena, of psychoacoustic effects resulting from the exploitation of machines/devices.

3. A way of modeling of vibroacoustic phenomena with the use of the signal parameterization method

Auditory sound impressions resulting from acoustic processes are strongly dependent on frequency, due to the physical conditions of acoustic wave propagation and perception. It should be regarded as significant, non-linearity and the range of perception of sound phenomena, in relation to the amplitude of sound pressure and frequency [2]. The author proposed a new approach to modeling vibroacoustic phenomena occurring in the environment of operated machines/devices. This approach relies on extracting selected features from the acoustic signal and from the audio signal. The acoustic signal is represented by the physical features of the sound. The audio signal contains information representing the subjective features of the sound. In the proposed method, the starting point for modeling vibroacoustic phenomena are the features of the acoustic signal and the features of the audio signal. Fig. 1 presents a method of modeling vibroacoustic phenomena occurring during the exploitation of machines/devices, which takes into account the psychoacoustic effects of noise.

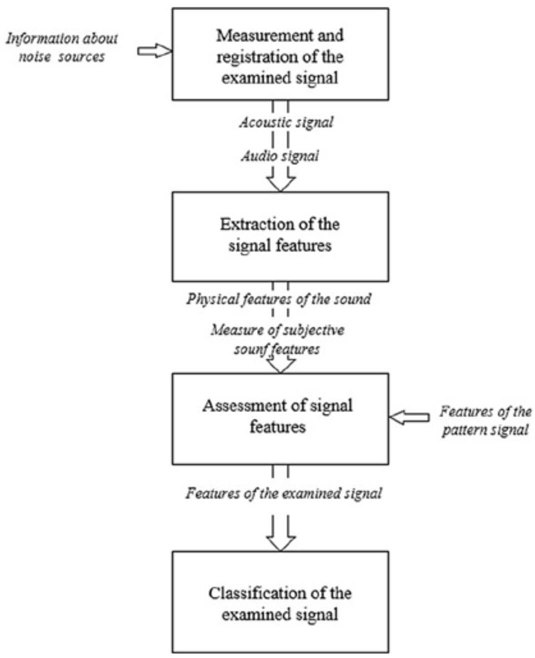


Fig. 1. The method of processing the features of the acoustic signal and the features of the audio signal

According to Fig. 1, the author proposed the following actions for the classification of the examined signal:

- measurement and recording of the examined signal: it is an information source for the acquisition and processing the features of the acoustic signal and the features of the audio signal,
- extraction of signal features: involves the use of a method/model for assessing physical and subjective features of an acoustic and audio signal,
- assessment of the features of the examined signal: the basis for the assessment of the features of the audio signal is the degree of compatibility of its features with the features of the reference signal,
- classification of the examined signal: the assignment of the audio signal to a given class of the reference signal based on the assessment of features.

The proposed method assumes, that assessment and selection of reference signals will be based on playing back recorded audio signals to recipients as part of a psychoacoustic experiment. Assignment of reference signals to the appropriate noise assessment class is realized on the basis of the results obtained from the signals presented during the experiment.

3.1. Description of the method of parameterizing the audio signal

In the considered research problem, the proposed method of parameterizing the audio signal is based on a perceptual model. This model uses the properties of the hearing mechanism of the human ear, characterized by a non-linear perception of the height of the frequency of received sound signals on a mel-scale. Method of parameterizing the audio signal, applied by the author, in modeling vibroacoustic phenomena is original and not found in the literature of the subject. The MFCC (Mel-Frequency Cepstral Coefficients) method is one of the most commonly used methods of speech signal parameterization. It allows the determination of a set of cepstral coefficients, i.e. the features of the signal from the melody spectrum [7]. Mel-cepstral coefficients are patterned on the processing of the acoustic signal in the cochlea of the human hearing organ. Their task is to reflect the natural response of the auditory system to sound stimulation. The proposed method consists in modeling the parameters extracted from the audio signal, which strongly depend on the subjective listening impressions of sound. The prerequisites for the implementation of the study of vibroacoustic phenomena by mel-cepstral coefficients was the recognition of the possibility of using the method of parameterizing the audio signal. This method allows taking into account, among others:

- randomness of signals commonly found in exploited machines/devices,
- variability of the frequency structure in the signal waveform,
- estimation of the signal spectra on a subjective perceptual scale,
- nonlinearity of the perception of the frequency of sound by human.

Determination of mel-cepstral coefficients gives the possibilities of effective classification and assessment of the examined audio signals. Lack of universality of solutions in the recognition of acoustic signal patterns for the purposes of their assessment is not due to the imperfection of methods but to the complexity of the source signals. For this reason, the transformation of the analyzed signals is used to obtain the appropriate space of sound features. For the purpose of

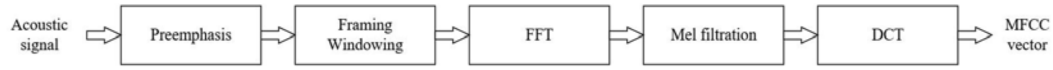


Fig. 2. Parameterization procedure of the audio signal using MFCC coefficients [25]

determining the vector of MFCC coefficients, it was applied an algorithm to extract the features of the audio signal (fig. 2).

According to the procedure presented in Fig. 2, it was adopted staged realization of signal features processing [25]:

Stage 1: Preemphasis process consisting of forming filtration, which results in weakening of low-frequency components and amplification of high-frequency components.

Stage 2: Signal framing, i.e. dividing the signal into short fragments called frames. It is possible applying overlapping successive time frames. Then, at this stage, windowing is carried out using Hamming window:

$$Ham(N) = 0,54 - 0,46\cos(2\pi\frac{n-1}{N-1}) \tag{1}$$

where:

N - frame length,
n=1,2,...,N.

Stage 3: Execution of the Fast Fourier Transform (FFT) algorithm on the windowed signal in individual frames and determination of the module of estimate of the power spectral density of the signal.

Stage 4: Performing mel-filtration using a set of bandpass triangular filters with frequencies determined in accordance with:

$$f_{mel} = 2595\log_{10}(1 + \frac{f_{\mathrm{Hz}}}{700}) \tag{2}$$

The calculations used the logarithm of energy, which reduces the sensitivity of filters to very loud and very quiet sounds and allows modeling of non-linear amplitude sensitivity of the human ear.

Stage 5: The final stage of the procedure is the use of the discrete cosine transform (DCT). The resulting vector MFCC coefficients is calculated according to the relationship:

$$MFCC_n = \sqrt{\frac{2}{N}}\sum_{i=1}^{N}\log(S_i)\cdot\cos\left[\frac{\pi n}{N}(i-0,5)\right] \tag{3}$$

$$S_i = \sum_{k=1}^{N}\left|X_r(k)\right|^2 H_i(k) \tag{4}$$

where:

i - filter number numer filtra,
X_r - frame spectrum,
H_i - filter set,
S_i - band energy,
n - coefficient number,
N - number of filters used.

In most recognition systems n is 1, and the coefficient $MFCC_0$ is omitted. The generated vector coefficients $\mathrm{MFCC_n}$ takes the form:

$$MFCC_n = \langle MFCC_1, MFCC_2, MFCC_3, ...MFCC_{13}\rangle \tag{5}$$

In the presented research method, it was assumed, that reference signals identified as part of the psychoacoustic experiment will be parameterized to determine mel-cepstral coefficients. According to the procedure (Fig. 2), mel-cepstral coefficients are determined for each examined audio signal, which will then be evaluated using pattern signals. Based on the determined MFCC coefficients, the classification of the examined signal consists in the assessment of the degree of compliance of its features with the pattern signals. In order to classify the examined audio signals (based on a set of standard signals), as a measure of assessment, the author proposed the smallest distance between two series of feature vectors, i.e. MFCC coefficients using the DTW method (Dynamic Time Warping). The signals emitted by vibroacoustic sources are characterized by dynamic variability in time, which causes that the features of these signals are also subject to variability. Dynamic Time Warping (DTW) is one of the methods used in speech recognition. In particular, this method was primarily used in recognizing isolated words and searching for passwords. This method is used to recognize and classify the matrices of features of the MFCC coefficients as a non-linear time transformation. It involves the transformation of the timeline, to better match two time sequences. Determining the distance using the Dynamic Time Warping (DTW) method consists in:

- calculating the so-called matrix of local distances d(m, n), which is created by calculating Euclidean distances between each vector of the examined signal and the pattern signal,
- determining the sum of local (Euclidean) distances, which is the distance accumulated along the optimal path in the local distance matrix, from the lower left corner of the matrix to its upper right corner.

Particular DTW distances between the pattern signals and examined signal can be determined according to the following formula:

$$d_{mn}(X,Y)=\sqrt{\sum_{k=1}^{K}(x_{k,m}-y_{k,n})\cdot(x_{k,m}-y_{k,n})} \tag{6}$$

where:

K- signal dimension,

m, n - vector sequences of X and Y features.

The accumulated distance calculated along the optimal path is the smallest of the possible accumulated distances of individual vectors of the MFCC coefficients of the examined signal and the pattern signal. It was assumed that the criterion for the classification of the examined signal in the scope of assessing noise annoyance will be the smallest DTW distance between mel-cepstral coefficients of the examined signal and a given pattern signal of the same time length.

3.2. Example of using the method and classification of the examined audio signal

The audio signal selected for examine was a sound sample called DW13a_9_56, characterized by an equivalent sound level L_{Aeq}=56[dB(A)]. The signal was recorded in a point located at the height of the road lane and it included the measurement of the immission of road noise sources of moving vehicles on a cobblestone pavement. Due to the research goal, the measurement was not made in accordance with the applicable methodology for carrying out environmental measurements [22]. The following set of devices and measuring aids were used for recordings and measurements:

- Brüel & Kjær 2238 Mediator sound level meter,
- Brüel & Kjær 4188 measuring microphone for measurements of a free field,
- ZOOM Handy Recorder H4n sound recorder,
- microphone tripod.

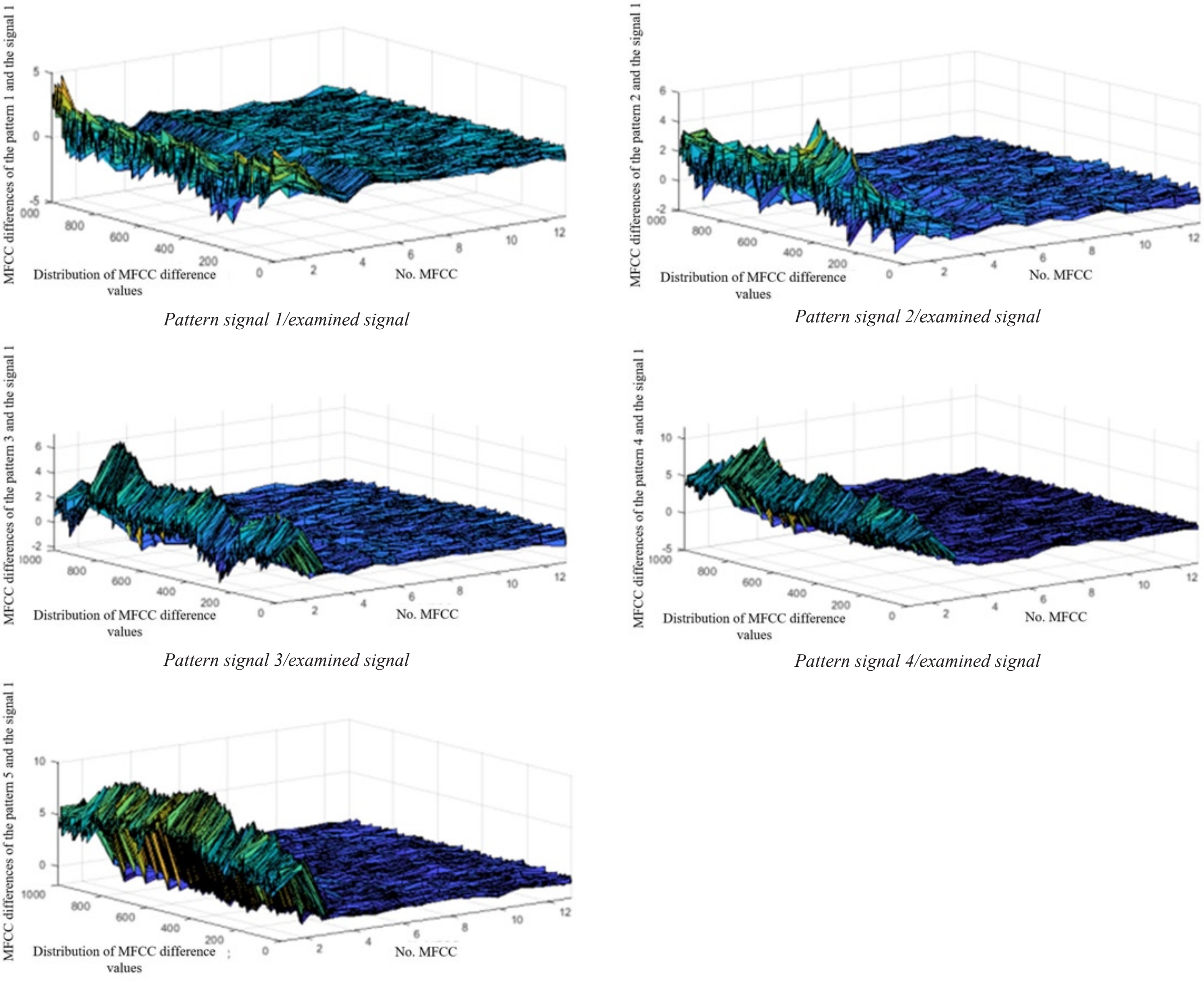
The time of recording and measurements was set at 5 [min]. During recording and measuring, the microphone was placed on a tripod in the middle of the pavement, at a height of 1.7 [m], which corresponded to the approximated height of the head of a potential passerby. It was assumed that in the subjective assessment of sound samples (in a psychoacoustic experiment), there participated people with normal hearing. Ten 10-second files were selected for each 5-minute file. In the carried out experiment, there attended 80 people aged between 22-50 years, who met the above assumption. In ongoing psychoacoustic research, to assess the annoyance of road noise from the 30 audio signals presented, a point scale in the range of 1 to 5 was proposed. Appropriate labels have been assigned to the assessed signals, i.e.:

- not at all - grade 1,
- little - grade 2,
- medium - grade 3,
- very - grade 4,
- intolerable - grade 5.

Taking into account the extent of the obtained sound level values corresponding to different measurement locations (characterized by variable buildings, variable road infrastructure system) and different road surfaces (i.e. pavement, asphalt), it was decided that the selected signal samples will be prepared in such a way that their equivalent

Table 1. List of σ and $\bar{x}$ values of MFCCn mel-cepstral coefficients for pattern signals

	Pattern signal 1		Pattern signal 2		Pattern signal 3		Pattern signal 4		Pattern signal 5	
Coefficient No.	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$
MFCC1	1,25	-2,07	1,40	-2,08	1,16	-1,99	1,38	0,61	1,40	1,43
MFCC2	1,27	-18,01	1,81	-16,32	1,62	-14,21	2,01	-11,14	2,00	-10,60
MFCC3	0,35	3,26	0,31	3,20	0,35	3,11	0,43	3,27	0,43	2,79
MFCC4	0,25	-0,51	0,25	-0,65	0,24	-0,99	0,33	-0,75	0,61	0,13
MFCC5	0,23	0,38	0,23	0,30	0,20	-0,09	0,28	-0,30	0,36	0,03
MFCC6	0,20	0,06	0,22	0,03	0,20	-0,03	0,31	-0,07	0,32	0,03
MFCC7	0,18	0,28	0,18	0,29	0,18	0,16	0,24	0,24	0,29	0,26
MFCC8	0,20	0,11	0,19	0,08	0,18	-0,09	0,21	0,24	0,32	0,19
MFCC9	0,18	0,13	0,20	0,16	0,18	0,13	0,21	0,47	0,34	0,18
MFCC10	0,17	0,02	0,18	0,03	0,17	0,10	0,20	0,02	0,37	0,14
MFCC11	0,16	0,06	0,17	-0,03	0,18	0,06	0,20	0,18	0,39	0,04
MFCC12	0,17	0,02	0,18	-0,02	0,17	-0,01	0,27	-0,22	0,42	-0,02
MFCC13	0,16	0,03	0,17	-0,03	0,18	0,02	0,18	0,20	0,44	-0,01



Pattern signal 1/examined signal

Pattern signal 2/examined signal

Pattern signal 3/examined signal

Pattern signal 4/examined signal

Pattern signal 5/examined signal

Fig. 3. Comparative analysis of MFCC coefficients of the examined signal with pattern signals

level corresponded to one of the following five values, i.e.: 56, 62, 68, 74, 80[dB(A)]. Input data for determining the MFCCn vector were audio pattern signals in the time domain, selected on the basis of a psychoacoustic experiment and representing the appropriate classes of noise annoyance. As a result of the procedure (Fig. 2), 13 mel-cepstral coefficients were calculated for each of the pattern signals and the examined signal. The calculations were carried out in the Matlab R2018a environment using the function (mfcc). Standard deviation (σ) and average ($\bar{x}$) values were determined for each of the coefficients for 1000 data (values were calculated with a step 0,01[s]), (Table 1).

The comparative analysis of the MFCC values of the examined signal with the pattern signals showed that the largest (local) relative differences between the MFCC coefficients occurred when the examined signal was combined with the pattern signal 4 (Fig. 3).

Individual DTW distances between the pattern signals and the examined signal were determined using the function (dtw) in the Mathlab R2018a environment. After substitution into (6), there have been determined DTW distances between MFCC coefficients for individual pattern signals and the examined audio signal DW13a_9_56.

Table 2 shows that the smallest accumulated DTW distance was obtained for the variant of assessment of the examined signal DW13a_9_56

Table 2. The summary cumulative DTW distances of the examined signal with the pattern signals

No.	Type of variant	Accumulated distance DTW
1.	Pattern signal 1/examined signal	197,829
2.	Pattern signal 2/examined signal	249,447
3.	Pattern signal 3/examined signal	304,864
4.	Pattern signal 4/examined signal	496,279
5.	Pattern signal 5/examined signal	537,361

together with the pattern 1, which means that the examined signal has been classified into the noise annoyance class: not at all.

4. Discussion of results and conclusions

The obtained results of the classification of the examined audio signal, based on the assessment of its features, i.e. MFCC coefficients, justify the application of the method of parameterization of the audio signal in modeling of vibroacoustic phenomena of machines / devices, taking into account the psychoacoustic effects of noise. The presented method can be directly used in the tasks of assessing sound sources, shaping employee health, or developing design and construction solutions for noise protection. The proposed method of parameterization of the audio signal expands and complements the applied energetic approach to modeling vibroacoustic phenomena of machines/devices. The possibilities of cepstral analysis based on the characteristics of human hearing (mel-scale), were an important arguments for the author to look for a new way of research in the field of modeling vibroacoustic phenomena. The presented method allows for classification of any audio signal in terms of noise assessment based on the mel-scale of hearing, for the identified source type of vibroacoustic energy. An important advantage of using the method of parameterizing the audio signal is the ability to model cepstrum, which allows you to include noise perception and interpretation of the information contained in the spectrum of the audio signal. The parameters determined in a perceptual scale reflect natural sound experiences, which is important for the assessment of psychoacoustic noise and vibration phenomena. The developed method of modeling vibroacoustic phenomena is based on the assessment of the acoustic/audio signal using the following methods:

- methods of scaling the audio signal (subjective assessment of acoustic impressions as part of a psychoacoustic experiment),
- methods of parameterizing the audio signal (assessment of signal features on the mel-scale), which is an objective description of perceived noise.

The undertaken research will be continued in the area of applicability of methods and models describing the psychoacoustic effects of noise in the time-frequency domain and other perceptual scale, e.g. the bark-scale. The method of modeling vibroacoustic phenomena proposed by the author gives new possibilities in the scope of machine/device assessment, at the stage:

- design and construction, as an additional criterion in the process of optimizing the silent operation of elements,
- exploitation, as part of diagnostic tests for the purposes of verifying the degree of wear of individual components and subassemblies, and, as a result, assessing their reliability level.

The processed information in the form of acquired features from the audio signal is an added value for the acoustic signal and can be effectively used in supporting decision making of the above-mentioned tasks.

Acknowledgement

The article is the result of research conducted in the Department of Production Engineering of the Faculty of Organisation and Management of the Silesian University of Technology, which was financed by the Rector's Habilitation Grant number 13/030/RGH19/0057.

References

1. Batko W, Dąbrowski Z, Engel Z, Kiciński J, Weyna S. Nowoczesne badania procesów wibroakustycznych. Radom: Wydawnictwo Instytutu Technologii Eksploatacji -PIB, 2005.
2. Cempel Cz. Wibroakustyka stosowana. Warszawa: Państwowe Wydawnictwo Naukowe, 1989.
3. Daniewski K, Kosicka E, Mazurkiewicz D. Analysis of the correctness of determination of the effectiveness of maintenance service actions. Management and Production Engineering Review 2018; 9(2): 20-25.
4. Directive 2002/49/Ec of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental Noise. Official Journal of the European Communities - 18.07.2002.
5. Engel Z, Piechowicz J, Pleban D, Strychniewicz L. Minimalizacja przemysłowych zagrożeń wibroakustycznych. Poradnik. Warszawa: Centralny Instytut Ochrony Pracy - Państwowy Instytut Badawczy, 2005.
6. Engel Z, Zawieska Wiktor M. Hałas i drgania w procesach pracy. Źródła, ocena, zagrożenia. Warszawa: Centralny Instytut Ochrony Pracy - Państwowy Instytut Badawczy, 2010.
7. Gauchev T, Fakotakis N, Kokkinakis G. Comparative evaluation of various mfcc implementations on the speaker verification tasks. Proc. of the SPECOM 2005: 191-194.
8. Genuit K. Fiebig A. Psychoacoustics and its benefit for the soundscape approach. Acta Acustica 2006; 92: 1-7.
9. 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.
10. Jasiulewicz-Kaczmarek M., Zywica 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.
11. 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.
12. Licklider JCR. Correlates of the auditory stimulus. [w:] Handbook of Experimental Psychology (red. Stevens S.S.). New York: Wiley, 1966.
13. Loska A, Moczulski W, Wyczółkowski R, Dąbrowski A. Integrated system of control and management of exploitation of water supply system. Diagnostyka 2016; 17(1): 65-74.
14. Loska A. Methodology of variant assessment of exploitation policy using numerical taxonomy tools. Management Systems in Production Engineering 2015; 2(18): 98-104.
15. Makarewicz R. Dźwięk w środowisku. Poznań: Ośrodek Wydawnictw Naukowych, 1994.
16. Ozimek E, Kocinski J, Kutzner D. et al. Speech intelligibility for different spatial configurations target speech and competing noise source in a horizontal and median plane. Speech Communication. 2013; (10): 1021-1032, https://doi.org/10.1016/j.specom.2013.06.009.
17. Paszkowski W. The assessment of acoustic effects of exploited road vehicles with the use of subjective features of sound. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(3): 522-529, https://doi.org/10.17531/ein.2019.3.19.
18. Paszkowski W, Kotus J, Poremski T, Kostek B. Evaluation of sound quality features on environmental noise effects - a case study applied to

road traffic noise. Metrology and Measurement Systems 2018; 25(3): 517-531.
19. Paszkowski W, Sobiech M. The modeling of the acoustic condition of urban environment using noise annoyance assessment. Environmental Modeling & Assessment 2019; 24: 319-330, https://doi.org/10.1007/s10666-018-9643-1.
20. Raimbault M, Lavandier C, Berengier M. Ambient sound assessments in or urban environments: field studies in two French cities. Appl. Acoust. 2003; 64: 1341-1256, https://doi.org/10.1016/S0003-682X(03)00061-6.
21. Rozporządzenie Ministra Środowiska z dnia 1 października 2012 r. zmieniające rozporządzenie w sprawie dopuszczalnych poziomów hałasu w środowisku, (Dz.U. z 8 października 2012, poz. 1109).
22. Rozporządzenie Ministra Środowiska z dnia 30 października 2014r. w sprawie wymagań w zakresie prowadzenia pomiarów wielkości emisji oraz pomiarów ilości pobieranej wody,Dz.U.2014, poz. 1542.
23. Rudno-Rudzińska B. Modelowanie emisji i propagacji dźwięku do prognozowania klimatu akustycznego środowiska zurbanizowanego. Prace naukowe Instytutu Telekomunikacji i Akustyki Nr 75. Monografie nr 39. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, 1994.
24. Schomer PD. Criteria for assessment of noise annoyance. Noise Control Eng. J. 2005; 53(4): 132-144, https://doi.org/10.3397/1.2839251.
25. Zieliński T. Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań. Warszawa: Wydawnictwa Komunikacji i Łączności WKŁ, 2007.
26. Zieliński T.P, Korohoda P, Rumian R. Cyfrowe przetwarzanie sygnałów w telekomunikacji. Warszawa: Wydawnictwo Naukowe PWN SA, 2014.

Waldemar PASZKOWSKI
Silesian University of Technology
Faculty of Organization and Management
ul. Roosevelta 26-28, 41-800 Zabrze, Poland

E-mail: wpaszkowski@polsl.pl

Article citation info:
LIANG H, MI J, BAI L, CHENG Y. Imprecise sensitivity analysis of system reliability based on the Bayesian network and probability box. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 508–519, http://dx.doi.org/10.17531/ein.2020.3.14.

He LIANG
Jinhua MI
Libing BAI
Yuhua CHENG

IMPRECISE SENSITIVITY ANALYSIS OF SYSTEM RELIABILITY BASED ON THE BAYESIAN NETWORK AND PROBABILITY BOX

NIEDOKŁADNA ANALIZA CZUŁOŚCIOWA NIEZAWODNOŚCI SYSTEMU W OPARCIU O SIEĆ BAYESOWSKĄ I POLE PRAWDOPODOBIEŃSTWA (P-BOX)

Sensitivity analysis measures how changes in system inputs affect outputs. Previously, a large amount of sensitivity analysis research was relevant to the precise probability that is regarded as an ideal condition of engineering. Due to insufficient test samples and the low accuracy of test data, system reliability with hybrid uncertainty is difficult to be described as a precise value. As a profusion of highly integrated electromechanical equipment is applied in modern life, it is impossible to apply sufficient resources to eliminate the stochastic property of every component, which necessitates the identification of highly sensitive components to efficiently reduce imprecision. Hence, based on the theory of imprecise probability, imprecise sensitivity analysis has become a popular research topic in the last decade. In this paper, a method for uncertain system reliability and imprecise sensitivity analysis is proposed based on a Bayesian network, a probability box and the pinching method. The feasibility and accuracy of the combined method are fully verified through the evaluation and analysis of a numerical example and a case study of an electromechanical system, and the highly sensitive components that heavily influence the imprecision of system outputs are accurately identified.

Keywords: *bayesian network; probability box; sensitivity analysis; reliability analysis.*

Celem analizy czułościowej jest badanie w jakim stopniu zmiany danych wejściowych systemu wpływają na dane wyjściowe. Dotychczasowe badania z wykorzystaniem analizy czułościowej były związane z dokładnym prawdopodobieństwem postrzeganym w inżynierii jako warunek idealny. Przy niewystarczającej wielkości badanej próby i niskiej dokładności danych testowych, niezawodność systemu o hybrydowej niepewności trudno opisać w sposób dokładny. Biorąc pod uwagę fakt, że we współczesnym świecie wykorzystuje się duże ilości wysoce zintegrowanych urządzeń elektromechanicznych, niemożliwa jest alokacja wystarczających zasobów w celu wyeliminowania właściwości stochastycznych każdego elementu. Oznacza to, że aby zredukować niedokładność, konieczna jest identyfikacja komponentów o wysokiej czułości. Dlatego też popularnym przedmiotem badań ostatniej dekady stała się niedokładna analiza czułości, bazująca na teorii niedokładnego prawdopodobieństwa. W artykule zaproponowano metodę analizy niezawodności niepewnego systemu jak również niedokładnej analizy czułościowej w oparciu o sieć bayesowską, pole prawdopodobieństwa i metodę pinch point. Możliwość wykorzystania i dokładność metody zostały w pełni potwierdzone na podstawie przykładu liczbowego jak również studium przypadku systemu elektromechanicznego; proponowana metoda pozwoliła na poprawne określenie wysoce czułych elementów systemu, które w dużym stopniu wpływają na niedokładność danych wyjściowych układu.

Słowa kluczowe: *sieć bayesowska; pole prawdopodobieństwa; analiza czułości; analiza niezawodności.*

1. Introduction

With the improvement of industrial techniques and requirements for productivity, plenty of high-integrity and complex-structured electromechanical systems (EMSs) have been widely employed and utilized. The performance of the equipment can be further enhanced with a higher integration, which necessitates a better understanding of the failure degradation law of key components. For the design of a highly integrated system, reliability theory has attracted considerable researches in recent years to study the failure relationship among systems and components with the lifetime of products as the main research object. Specifically, many strategies have been proposed for the sake of reliability model establishment, such as block diagrams [1], Markov analysis (MA) [2], simplified equations [3] and fault trees (FTs) [4]. An FT is a powerful tool for reliability modeling that uses binary decision diagrams (BDDs). As an extension of an FT, a Bayesian network (BN) describes the relationship of failure events with a directed acyclic graph (DAG) as well as conditional probability tables (CPTs), as proposed by Pearl [5], achieving significant development in system reliability and safety analyses. In addition, Cai et al. [6] evaluated the reliability of a blowout preventer control system with a BN. A BN model was also established for wind turbines by Su et al. [7] to achieve a reliability analysis considering environmental factors and uncertainty. Mi et al. [8] presented a methodology to quantify the importance of common cause failures in the context of a BN and probability bounds analysis.

The uncertainty of the system is very important for the accuracy of the reliability estimation since it is difficult to attain a comprehensive knowledge of system failure. Specifically, in simulation and experimental processes, according to Ref. [9], uncertainties can be divided into three sources:

1) Uncertainties in parameterization.
2) Uncertainties in modeling.
3) Uncertainties in experiments.

However, to reduce the effects of uncertainty, it is more advantageous to take the intuitive uncertainty quantification metrics and the adjustment of the reliability analysis into prior consideration.

Uncertainty is currently divided into two types: epistemic (reducible) uncertainty and aleatory (irreducible) uncertainty [10, 11]. Aleatory uncertainty, determined by the random properties of a system, cannot be reduced, whereas the probability distribution can be derived easily by classic probability theory. However, epistemic uncertainty caused by the lack of knowledge of system mechanisms and samples cannot be eliminated by classic probability methods. Although the effects of epistemic uncertainty can be diminished through mass testing data and a deep understanding of system mechanisms, experts have not reached a consensus on dealing with epistemic uncertainty at present except for taking its quantification under initial consideration. The theory of evidence was first proposed by Dempster and then further promoted by Shafter, so it is called D-S evidence theory [12]. Basically, it can be interpreted as a generalization of Bayesian probability, assigning a number between 0 and 1 to the degree of belief supporting a certain proposal [13]. The details of the definitions refer to references [13–15]. Miscuri et al. [13] utilized an evidence network, which is the combination of the BN and evidence theory, and critical networks for security vulnerability assessment. In addition to evidence theory, probability bounds theory (PBA), also known as the probability box (p-box), is another popular uncertainty quantification metric. Based on precise probability theory, the p-box is divided into parametric and nonparametric types. The parametric p-box assumes that the probability distributions of the variables are known, and the possible cumulative distribution functions (CDFs) of the variable are in the same distribution. However, for the nonparametric p-box, the CDFs can be any CDF between the lower and upper probability bounds. Mi et al. [16] constructed a p-box to characterize the uncertainty of a multistate system with CCF. Feng et al. [17] evaluated sensitivity by utilizing the p-box as the quantification metric and a survival signature as the reliability modeling method. Meanwhile, Schöbi et al. [18] proposed interval-valued Sobol indices as an extension of classic definition by modeling the uncertain input parameters through parametric p-boxes. In short, the p-box is suitable for illustrating the epistemic uncertainty caused by insufficient samples, while it is more beneficial to consider evidence theory for the uncertainty caused by low data accuracy [19]. As Ref. [20] concludes, for any event $U \in F$ ， the upper and lower probability bounds respectively correspond to the belief function $Bel(U)$ and plausibility function $Pl(U)$, in which we can find the mutual conversion of evidence theory and p-box in the mathematical form.

Sensitivity analysis (SA) quantifies the influence of input uncertainty variation on the system output uncertainty. The purpose is to determine the main source of the system uncertainties. SA provides a basis for uncertainty reduction and can improve the robustness of the model prediction. Traditionally, SA methods for precise probability distribution have been developed rapidly, and various approaches have been proposed, such as regional sensitivity analysis [21] and matrix-based metrics [22].

However, there are still few publications for imprecise sensitivity analysis (ISA) [18]. Sankararaman & Mahadevan [23] and Krzykacz-Hausmann [24] described a global SA in the presence of Bayesian hierarchical models. Ref. [25] introduced Sobol indices for ISA. In addition, Helton et al. [26], on the basis of evidence theory, discussed the variance-based algorithm. The pinching method, proposed by Ferson [27], compares the variation in output uncertainty when part of the input variables has eliminated the uncertainty as a precise value, interval or probability distribution.

In response to the necessity of ISA studies of uncertain system reliability, this paper proposes a method to establish a reliability model with a BN, using the pinching method [27] to complete sensitivity analysis with the imprecision characterized by the p-box. Then, the high-sensitivity components and subsystems can be identified by ranking the indices. This approach is introduced as a new solution that implements the Bayesian network and pinching method for reliability and sensitivity analysis. A numerical example and an EMS case are detailed and analyzed by the proposed ISA approach to verify its feasibility. Hence, this article is organized as follows. Section 2 introduces the reliability, uncertainty, and sensitivity analysis theories involved in the following cases. Using an example of an uncertain system, details of the reliability modeling and sensitivity analysis are described in Section 3. In Section 4, the proposed method is applied to an EMS. Conclusions are provided in Section 5.

2. Preliminaries

2.1. Bayesian network

A BN consists of a DAG and CPTs, representing the direct dependency probability relationships among the variables [28]. Fig. 1 shows a simple BN, where the texts in the circles refer to certain failure events, and the directed arrows indicate the relationship of events. In the graph, nodes with only outputs are named root nodes, whereas leaf nodes have only inputs. Therefore, the clear and brief form to illustrate the propagation of failures is the advantage of a DAG. For constructing CPTs in the reliability and safety field, when the logic relation of parent nodes is AND, it means that the child event could be true only if parent events are true. Moreover, if the logic relation is OR, the child event will be true as long as one parent event is true. Notably, to optimize the calculation, CPTs should follow some format specifications. In this paper, "F" refers to the fail state of the component and "T" refers to the normal state. As a proposition regarding whether the given component state is true in the CPT table, "1" means true, and "0" means false. It is assumed that X_1 is in series with X_2 and that X_3 is in parallel with Y. As shown in Table 1 and Table 2, the tables that describe the marginal probability distribution of Y and T are CPTs, and the two tables depict the logic relation of AND and OR, respectively.

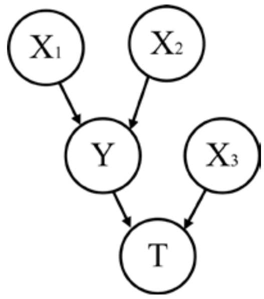


Fig. 1. A simple BN

Table 1. CPT of intermediate node Y

X_1	X_2	Y	
		F	T
F	F	1	0
F	T	0	1
T	F	0	1
T	T	0	1

The reasoning of the BN consists of forward and backward inference, also termed as predictive and diagnostic analysis, respectively. The former infers the marginal probability of any node in the condition of a given parent node's prior marginal probability mass function (PMF) and conditional PMFs of other child nodes in the network.

Table 2. CPT of leaf node T

X_1	X_2	T	
		F	T
F	F	1	0
F	T	1	0
T	F	1	0
T	T	0	1

Assuming that a set of random variables $\{X_1,...,X_n\}$ is composed of the system failure events corresponding to each node in a BN, the joint distribution can be calculated by the following formula:

$$P\{X_1,...,X_n\}=\prod_{i=1}^{n}P(X_i|\pi_i) \tag{1}$$

where π_i is the parent node of X_i. For the BN shown in Fig. 1, the joint probability distribution is given by Eq. (2):

$$P\{T,Y,X_1,X_2,X_3\}=P(T|X_3,Y)P(Y|X_1,X_2)P(X_3)P(X_2)P(X_1) \tag{2}$$

Using Eq. (1), the marginal probability distribution of X_i can be presented as:

$$P\{X_i\}=\sum_{except\ X_i}P\{X_1,...,X_n\} \tag{3}$$

With R defined as the reliability of the system, the reliability of the system shown in Fig. 1 can be presented as:

$$R=1-\sum_{except\ T}P\{X_1,X_2,X_3,\text{Y},\text{T}\} \tag{4}$$

2.2. Probability box

For a system with aleatory uncertainty, precise probability distributions can be used to quantify the degree of uncertainty, such as exponential, Weibull, and lognormal distributions. Consequently, classic probability theory exhibits favorable performance for quantitative issues of aleatory uncertainty. However, due to the incomplete knowledge of the system mechanism and the sample data, epistemic uncertainty always exists in the system. Furthermore, the classic probability method is not the appropriate evaluation approach because of the probability parameters defined as the intervals. To precisely measure the system uncertainty, the p-box is a solution providing a clear view of the epistemic uncertainty of a random variable and has been widely applied to quantify and represent the uncertainty in risk analysis [29] [10]. A nonnegative random variable X describes the lifetime of a component. $F_L(t)$ and $F^U(t)$ are CDFs of variable X on real number R, and $F(t)=P\{X\le t\}$. Suppose F is a set of nondecreasing functions that map R into [0,1], where $F_L(t)$ and $F^U(t)$ are the lower and upper bounds of F. Then, a p-box is defined by a probability family that matches the constraints $F_L(t)\le F(t)\le F^U(t)$ and $F(t)\in$F [10]. For reliability $R(t)$=1-$F(t)$, the p-box $\mathfrak{R}$ reflects the survival probability and is defined as:

$$\mathfrak{R}=\left\{R(t),\forall t\in R\,\middle|\,R_L(t)\le R(t)\le R^U(t)\right\} \tag{5}$$

For example, suppose the random variable X_{wb} follows a Weibull distribution, the shape and scale parameters are set as β=3 and η=[10,40], respectively, and the parameters for the lognormal distribution X_{Logn} are σ=0.4 and μ=[6,8]. As shown in Fig. 2, the p-boxes describe the reliability bounds by $R_L(X_{wb})$, $R^U(X_{wb})$, $R_L(X_{Logn})$ and $R^U(X_{Logn})$, and the uncertainty can be quantified as the regions of S_{wb} and S_{Logn}. The area of epistemic uncertainty space enclosed by the upper and lower bounds can be converted from graphs to numerical form via Eq. (6):

$$S_s=\int_0^{+\infty}(1-F_L(t))dt-\int_0^{+\infty}(1-F^U(t))dt=\int_0^{+\infty}R^U(t)dt-\int_0^{+\infty}R_L(t)dt=ET^U-ET_L. \tag{6}$$

where ET^U and ET_L represent the maximum and minimum mean lifetimes of the system, respectively. Eq. (6) quantifies the variation of system reliability with epistemic uncertainty, providing an index for uncertainty reduction. Moreover, the index is calculated for SA in the next subsection.

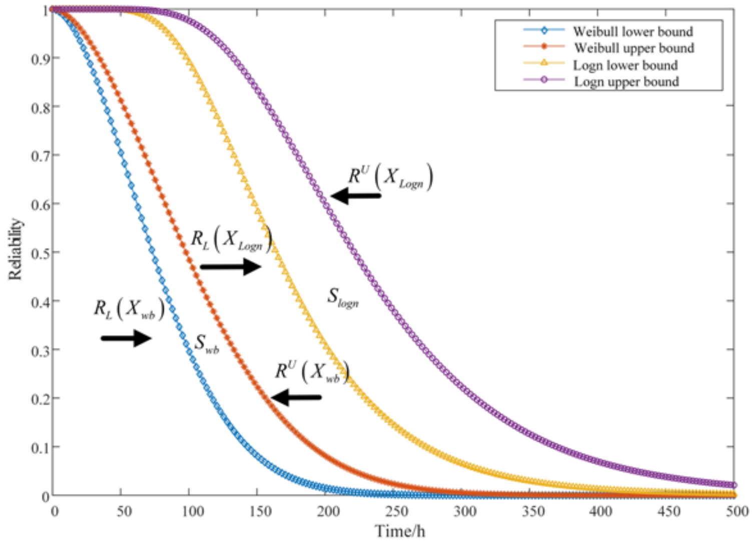


Fig. 2. P-boxes of Weibull and lognormal distributions

2.3. Sensitivity analysis of reliability

It has been shown that increasing the quantity and improving the accuracy of samples can reduce the epistemic uncertainty, but this is difficult to achieve for every component in a complex system. System reliability SA [30] [26] identifies the high-sensitivity components and optimizes their uncertainty properties to enhance equipment performance and save resources. Therefore, a variety of SA metrics have been developed for better performance to solve practical engineering issues.

In engineering practice, since the precise distribution is always unknown, an interval value is used for uncertainty prediction, and traditional SA methods are invalid due to the imprecise form of input variables. Ferson [27] proposed the p-box to characterize this uncertainty, which merges interval analysis and classic probability theory and treats aleatory and epistemic uncertainty separately on the basis of maintaining their features. The p-box permits a comprehensive uncertainty analysis, and this fact obviates some of the complexity that afflicts traditional Monte Carlo approaches to sensitivity analysis based on similar ideas. Based on the uncertainty definition by the p-box, the sensitivity index S_e can be computed by Eq. (7):

$$S_e = 1 - \frac{un(\mathrm{A})}{un(B)} \tag{7}$$

where B is the initial value of the epistemic uncertainty and A is the uncertainty index when the input epistemic uncertainty is reduced. Moreover, $un()$ represents the uncertainty quantification method, which can be described by the p-box graph and calculated through the size of the uncertainty space enclosed by the probability bound; the details can be seen in Section 2.2.

To identify highly sensitive components, the pinching method should be initially used to reduce the uncertainty of each component, followed by evaluating and ranking the sensitivity indexes obtained by Eq. (7). It should be noted that, unlike the variance-based index, the uncertainty reduction will not add up to 100% after all the input variables eliminate the uncertainty.

There are multiple possibilities to pinch uncertainty. Different pinching strategies provide diverse sensitivity values but will not affect the ranking of the values. The three different strategies are listed as follows [27]:

(i) replace an input with a point value,

(ii) replace an input with a precise distribution function, or

(iii) replace an input with a zero-variance interval.

Considering that aleatory uncertainty is easy to model by classic probability theory but hard to eliminate, this work selects strategy (ii) to pinch the uncertainty, which focuses on the characterization of aleatory uncertainty and the elimination of epistemic uncertainty. Due to the unknown target parameter value after pinching uncertainty, different target values used in the SA cause changes in index values, which is termed the deviation. To reduce the effect of deviation, a new sensitivity index that takes the means of all sensitivity index values is proposed and is shown as Eq. (8):

$$S_{mean} = \frac{1}{l_s} \sum_{i=0}^{l_s} 1 - \frac{un(\mathrm{A}_i)}{un(B)} \tag{8}$$

where l_s is the sampling value of the random variable parameter.

3. Reliability sensitivity analysis for an uncertain system

3.1. Analysis process

In this section, a comprehensive work for reliability sensitivity analysis is presented, and there are 3 steps for the ISA of system reliability, as shown in Fig. 3.

i Preprocessing: The aim of this step is to determine the distribution of pinched variables and construct the input vectors. Furthermore, based on the failure mechanism of the system, the BN model needs to be prepared for forward inference to evaluate system reliability.

ii Uncertain system reliability analysis: Via the inference of the BN, the probability distribution families of system reliability in the case of pinched input variables and initial inputs are obtained. Next, a p-box is employed to quantify the uncertainty and visualize it. The uncertainty quantification index is defined as the size of the region enclosed by probability bounds.

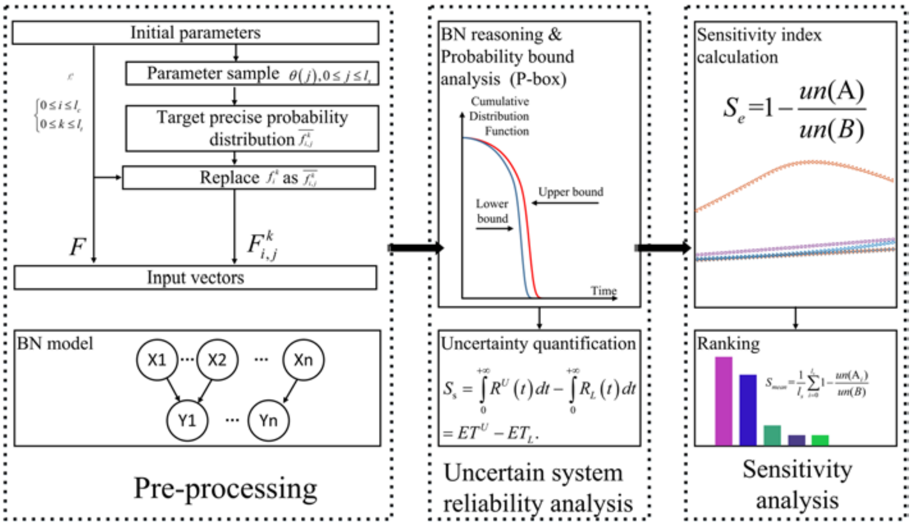


Fig. 3. The analysis process of this work

iii Sensitivity analysis: Because of the deviation of the sensitivity, computation of the mean of sensitivity values as the ultimate index of sensitivity assessment is applied to reduce the effect of deviation, which briefly indicates the highly sensitive components in the system.

The detailed analysis process will be described in the following sections.

3.2. Preprocessing

3.2.1. Input vectors

Assume that a system consists of l_c components in l_t different types, where the index of a component is defined as i, k refers to the type index, and that they match the constraints $k \in \{1,2,\ldots,l_t\}$ and $i \in \{1,2,\ldots,l_c\}$. For example, f_i^k expresses the failure probability of component i with type k. The input vector in the condition that no variable is pinched is written as the initial vector $F = \left[f_1^1, \ldots, f_i^k, \ldots, f_{l_c}^{l_t} \right]$. Then, as Section 2.3 describes, when f_i^k is pinched, the epistemic uncertainty of the i-th component with type k. is hypothetically eliminated. Hence, suppose $\overline{f_i^k}$ is the distribution f_i^k after pinching. Moreover, f_i^k will be replaced by $\overline{f_i^k}$ in the initial input vector. The input vector F will be renewed and written as F_i^k, where $F_i^k = \left[f_1^1, \ldots, \overline{f_i^k}, \ldots, f_{l_c}^{l_t} \right]$. For instance, in the case of f_5^2 being pinched, the F_5^2 matches the equation $F_5^2 = \left[f_1^1, \ldots, \overline{f_5^2}, \ldots, f_{13}^5 \right]$. Because the target of pinching is just a hypothesis, diverse target parameters lead to different input vectors, obviously affecting the ISA results, which is called deviation. Therefore, $\overline{f_i^k}$ is sampled with sample value l_s for comprehensive analysis results and defined by target parameter $\theta(j)\left(0 < j \le l_s\right)$ as $\overline{f_{i,j}^k}$. Similarly, f_i^k is replaced by $\overline{f_{i,j}^k}$ for an input vector, where $F_{i,j}^k = \left[f_{1,j}^1, \ldots, \overline{f_{i,j}^k}, \ldots, f_{l_c,j}^{l_t} \right]$. Similar to the above example, when $l_s = 10000$ and the target parameter is $\theta(400)$, the input vector is written as $F_{5,400}^2 = \left[f_1^1, \ldots, \overline{f_{5,400}^2}, \ldots, f_{13}^5 \right]$.

3.2.2 BN modeling

According to Section 3.2.1, the input vector $F_{i,j}^k$ when the probability distribution of component i is pinched, uncertainty is obtained.

Subsequently, the following matrix D can be used to illustrate the relationships of each DAG node shown as Eq. (9):

$$D=\begin{bmatrix} d_{1,dm} & \cdots & d_{1,dm} \\ \vdots & \ddots & \vdots \\ d_{1,dm} & \cdots & d_{22,dm} \end{bmatrix} \tag{9}$$

where $d_{a,}$b=1 means the arrow in the DAG moves from node a to node b. When there is no connection between a and b, $d_{a,c}=0$. After matrix D is obtained from the BN model, referring to Table 1 and Table 2, the CPTs of the child nodes in the DAG can be listed in the form of a column vector to participate in the forward inference.

3.3 Uncertainty system reliability analysis

Combining Eq. (3) with Eq. (1), the marginal probability distribution of child nodes can be written as:

$$P\{\mathrm{n}\}=\mathrm{CPT}_{\mathrm{n}}\prod_{i=p}P(\pi_i) \tag{10}$$

where n is the child node, π_i is a parent node, and p is the number of parent nodes. Therefore, $P\{\mathrm{n}\}$ represents the marginal probability of the failure events of child nodes. Similarly, $P\{\pi_i\}$ are the marginal probabilities for parent nodes. This formula accomplishes BN forward inference and deduces the system reliability.

Eq. (10) can be used to compute marginal probability when input variables are precise probability distributions. However, uncertainty exists in the parameters of $F_{i,j}^k$, and it is necessary to sample the parameters. Assume the sample value for BN inference is l_{bn}. Consequently, the input vector after sampling is $F_{i,j,m}^k=\left[f_{1,j,m}^1,\dots,\bar{f}_{i,j,m}^k,\dots,f_{l_c,j,m}^{l_t}\right]$. $P_{i,j,m}^k\{n\}$ can be written as:

$$P_{i,j,m}^k\{n\}=CPT_n\times\prod_{i=n_k}\left(F_{i,j,m}^k\times D_n\right) \tag{11}$$

where $P_{i,j,m}^k\{n\}$ is the probability of the event represented by the child node n . $F_{i,j,m}^k$ should add zeros to expand the size and be assigned during the iterations. The reliability can be written as:

$$R_{i,j,m}^k=1-P_{i,j,m}^k\{A\} \tag{12}$$

3.4. Sensitivity analysis

The pinching method is applied to analyze the sensitivity by eliminating the epistemic uncertainty of a variable and computing the change in the output uncertainty. To overcome the deviation issue of determining different target parameters *θ(j)*, the size of the area that is enclosed by the upper and lower bounds of the p-box should be calculated as the uncertainty quantification index. Then, based on the computed uncertainty index, the sensitivity index can be obtained by definition in Eq. (8).

3.5. Numerical example

In this section, a complex nonrepairable system from Ref. [31], composed of thirteen components with five different types, is described to demonstrate the effectiveness of the proposed method. Note that numbers in the solid line and the lower right corner denote the type of the component and the serial number, respectively, while Roman numerals and English letters represent subsystems. Table 3 gives the type of probability distributions and parameter ranges of each component, where η and β are the scale parameter with the hour unit and the nondimensional shape parameter of the Weibull distribution, respectively. Meanwhile, the λ of the exponential distribution represents its mean value with the same unit as η.

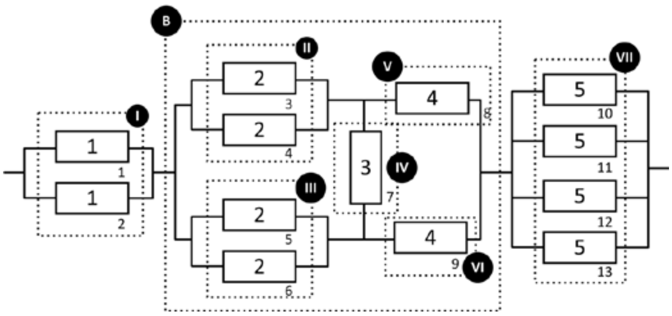


Fig. 4. A block diagram of a nonrepairable system [31]

Table 3. Probability distributions with epistemic and aleatory uncertainty of 5 types of components

Type	Distribution	Parameter (with epistemic uncertainty)
1	Weibull	$\eta_1=[1.68,1.86],\beta_1=2.08$
2	Exponential	$\lambda_2=[1.07,1.33]$
3	Weibull	$\eta_3=[2.12,2.51],\beta_3=1.38$
4	Weibull	$\eta_4=[2.99,3.41],\beta_4=2.51$
5	Exponential	$\lambda_5=[2.01,2.28]$

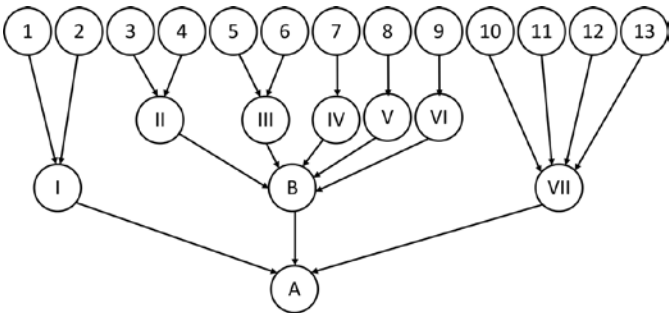


Fig. 5. BN of the system

Based on the definitions described above, the DAG shown in Fig. 5 could be transformed into the following matrix by Eq. (13) and Eq. (14).

$$D=\begin{bmatrix} d_{1,1} & \cdots & d_{1,22} \\ \vdots & \ddots & \vdots \\ d_{1,22} & \cdots & d_{22,22} \end{bmatrix} \tag{13}$$

$$
\text{where} \quad \begin{cases} d_{1,14} = d_{2,14} = 1 \\ d_{3,15} = d_{4,15} = 1 \\ d_{5,16} = d_{6,16} = 1 \\ d_{7,17} = 1 \\ d_{8,18} = d_{9,19} = 1 \\ d_{10,20} = d_{11,20} = d_{12,20} = d_{13,20} = 1 \\ d_{15,21} = d_{16,21} = d_{17,21} = d_{18,21} = 1 \\ d_{14,22} = d_{21,22} = d_{20,22} = 1 \\ \text{else} \ = 0 \end{cases} \tag{14}
$$

Furthermore, the CPT of each intermediate node in Fig. 5 can be established according to Section 3.3. For example, Table 4 shows the CPT of leaf node A based on the standards and specifications detailed in Section 2.1.

Note that it is more advantageous to express CPTs in a matrix form to simplify calculations. However, since subsystems II, III, IV, V, VI, and VII are not logical AND gates or OR gates, the CPT of node B must be listed separately. As shown in Table 5, CPT_B is a column vector with a length of $2^6 \times 1$.

Table 4. CPT of subsystem A

1	2	A	
		F	T
F	F	1	0
F	T	1	0
T	F	1	0
T	T	0	1

Table 5. CPT of subsystem B

II	III	IV	V	VI	B	
					F	T
F	F	F	F	F	1	0
T	F	F	F	F	1	0
F	T	F	F	F	1	0
...	...	...	...	...	...	...
T	F	T	T	T	0	1
F	T	T	T	T	0	1
T	T	T	T	T	0	1

Fig. 6 depicts the reliability p-box of each type component under different input vectors in the system according to the method described in Section 3.3. The impact of the components with the same connection methods, as well as failure probability distributions, can be considered equivalent. Therefore, simplification should be taken into prior consideration in reliability modeling. From the results shown in Fig. 7, we find that the probability bounds of the system reliability obtained by the uncertainty pinching of the same type of components completely overlap. Thus, we can conclude that for sensitivity analysis, the components with the same type and connection are of equivalence. Hence, only one component for each type needs to be selected to perform SA.

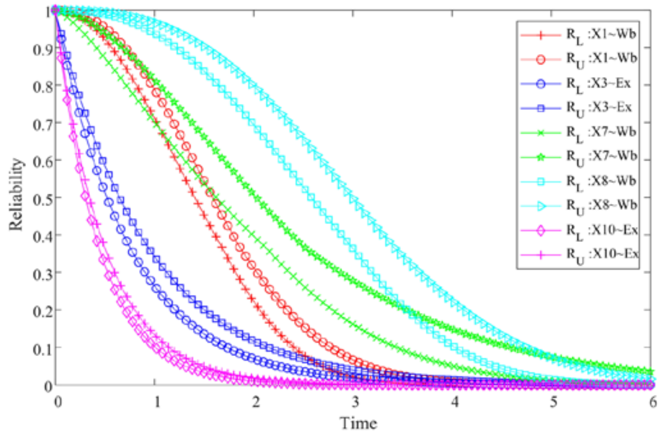


Fig. 6. Reliability p-box of different component types (X1, X3, X7, X8 and X10 refer to component types 1-5, respectively).

Fig. 8 depicts the p-box of the system reliability when the uncertainty of the random variables X1, X3, X7, X9 and X10 are pinched. Compared with the initial bounds, the bounds after pinching are slightly contracted. However, there may exist crossover and overlapping parts for the curves shown in Fig. 8, and it is difficult to compare the uncertainty space of p-boxes directly. Hence, it is more beneficial to define an index to quantify the uncertainty space of the system to further quantify the sensitivity. Therefore, measuring the size of the upper and lower areas is an appropriate method for uncertainty quantification. The ISA approach was described in Section 3.4.

With different values of the target parameter $\theta(j)$, the sensitivity of the component will change differently, which is denoted as the deviation. Fig. 9 shows the sensitivity deviation of components with different types. It is clear that the sensitivity of the type 5 components is significantly higher than others, which means that the uncertainty reduction in type 5 components will cause the greatest influence on the uncertainty of system reliability. However, the sensitivity of type 3 and type 4 components are much lower than other types, and it is hard to identify the type with the lowest sensitivity due to the sensitivity variation. According to Section 2.3, the influence of deviation can be reduced by the proposed method by calculating the mean value with Eq. (8). Fig. 10 shows a bar plot of the sensitivity of the components, and it can be clearly observed that the type 5 components are the most sensitive. Conversely, type 3 component X7 is the most insensitive component.

4. Case study

4.1. Description of the case

In Section 3, a comprehensive method is proposed for analyzing the reliability and sensitivity of the system, and a numerical example is introduced to detail the steps of the method. However, in a practical engineering system, the degradation of the components directly causes the working efficiency reduction, and the lack of sample data will also lead to the existence of uncertainty and nonlinear characteristics in a system. In this section, the proposed method is used to analyze the reliability and sensitivity of the electromechanical system in Ref. [10]. Fig. 11 shows the schematic of this electromechanical system, which is composed of a control system, a power supply system, a powertrain system, and a hydraulic system. More specifically, the control system includes two control modules connected in parallel to perform the start-stop control of the main valve and control execu-

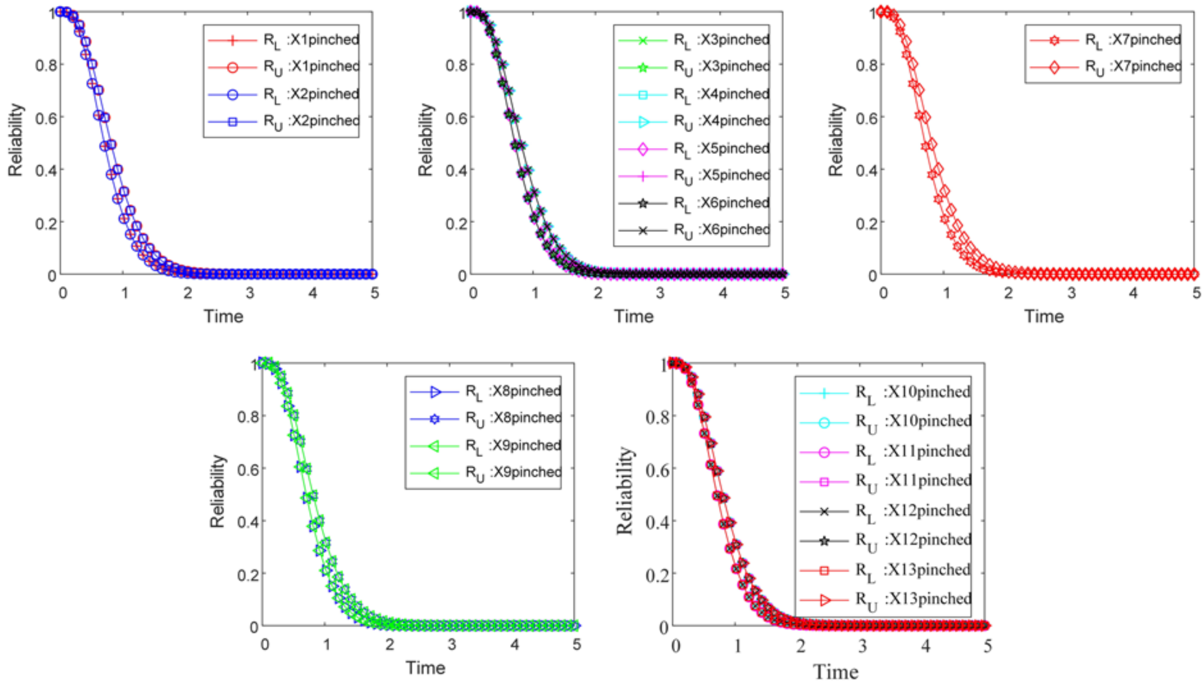


Fig. 7. Reliability p-boxes of components

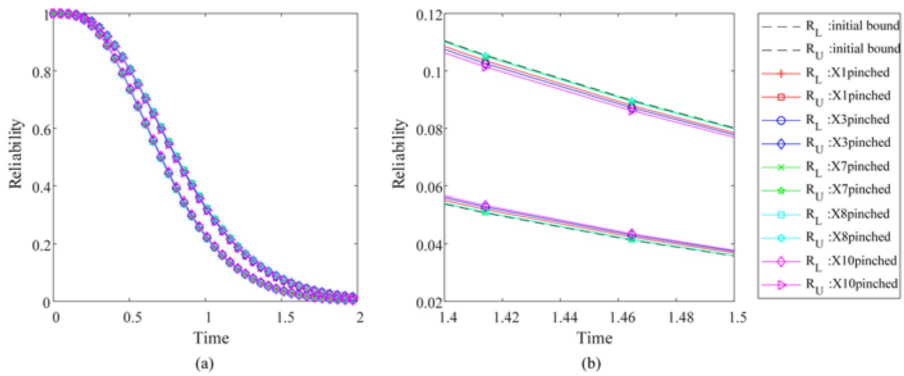


Fig. 8. P-boxes with diverse inputs pinched ((b) is the time interval [1.4, 1.5] of (a))

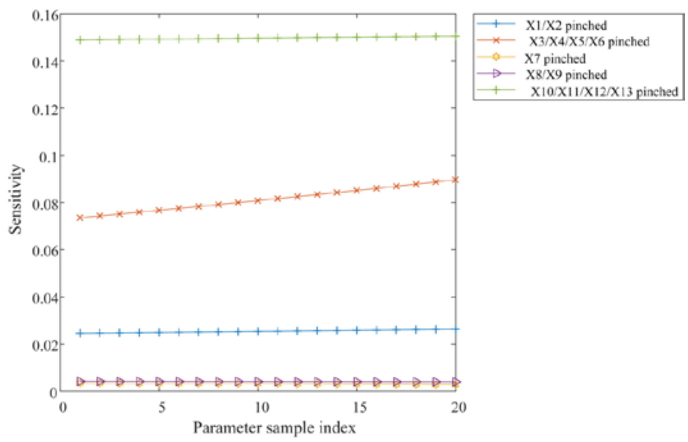


Fig. 9. Deviation of sensitivity

tion of the hydraulic subsystem. Meanwhile, the powertrain system is a key subsystem, which includes a turbine, a reducer and a pump. For the power supply subsystem, two valves are included in the emergency work mode, while only one main valve is contained in the main working mode. Based on the relationships among the components, the fault tree of the system can be plotted as shown in Fig. 12.

To introduce the method and simplify the calculations, the following assumptions are made for system reliability modeling:

1) A component or subsystem has the same failure probability distribution as its corresponding assembly component.

2) Components and subsystems whose failures rarely occur or do not cause system failure are negligible.

Assume that the failure probability of the basic components follows the Weibull distribution and lognormal distribution, respectively, according to the mechanical and electrical characteristics of the system. Additionally, based on accelerated life

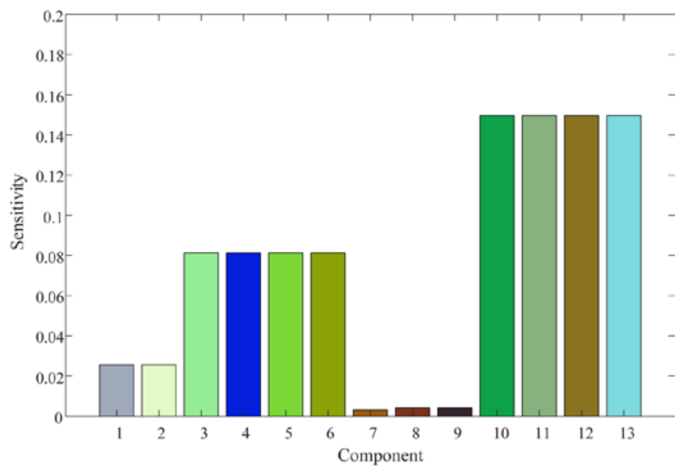


Fig. 10. Mean sensitivity of 13 components

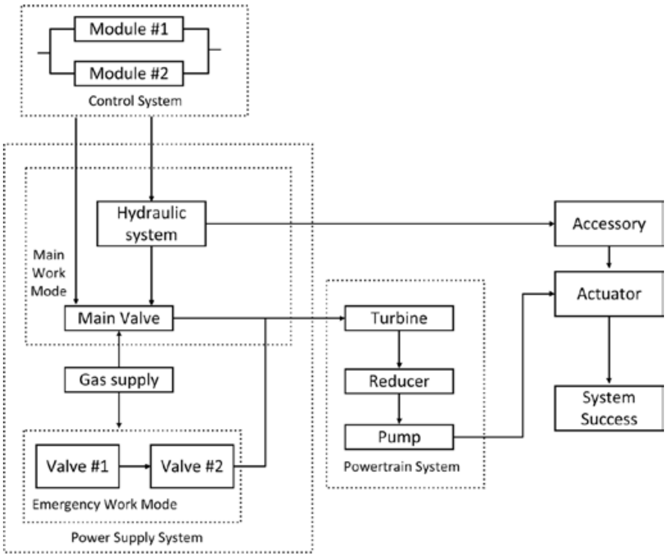


Fig. 11. Functional block diagram of an electromechanical control system

Table 6. Number and description of the events in the system

No.	Event description	No.	Event description
S	Complex electromechanical system task failure	X_3	Turbine failure
Y_1	Control system failure	X_4	Reducer failure
Y_2	Powertrain system failure	X_5	Pump failure
Y_3	Power is not transmitted to the subordinate unit	X_6	Valve #1 failure
Y_4	Main work mode failure	X_7	Valve #2 failure
Y_5	Emergency work mode failure	X_8	Main valve failure
X_1	Control module #1 failure	X_9	Hydraulic system failure
X_2	Control module #2 failure		

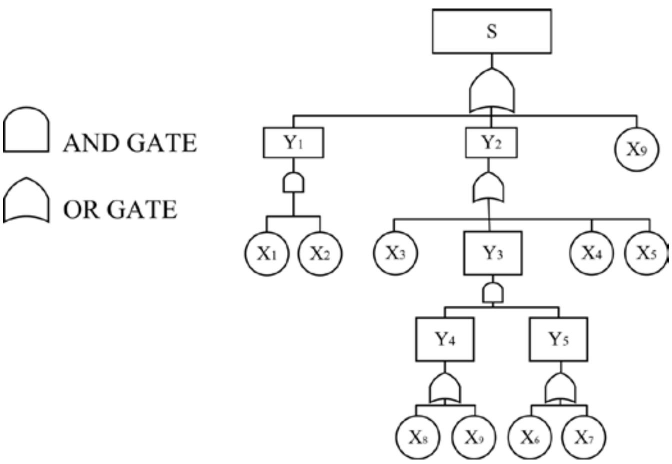


Fig. 12. Fault tree of the electromechanical control system

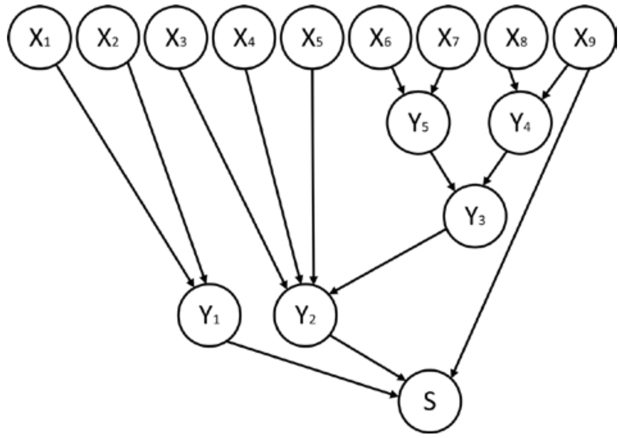


Fig. 13. BN of the electromechanical control system

testing and field data analysis, Table 7 lists the life distribution and life interval of different subsystems and components of the above-mentioned system [10].

4.2. System reliability modeling

Since it is necessary to obtain both the DAG and CPTs for the construction of a BN, the DAG can first be obtained and shown as based on the fault tree in Fig. 12. Meanwhile, owing to the forward reasoning requirement of a BN, it is essential to provide the matrix form of the DAG denoted as D, where the element $d_{i,j}$ in the matrix can be represented by Eq. (15).

$$\begin{cases} d_{1,10} = d_{2,10} = 1 \\ d_{3,11} = d_{4,11} = d_{5,11} = d_{12,11} = 1 \\ d_{6,14} = d_{7,14} = 1 \\ d_{8,13} = d_{9,13} = 1 \\ d_{13,12} = d_{14,12} = 1 \\ d_{10,15} = d_{11,15} = d_{9,15} = 1 \\ \text{else } = 0 \end{cases} \tag{15}$$

Additionally, the CPTs can be represented as the form of matrices CPT_{Y_1}, CPT_{Y_2}, CPT_{Y_3}, CPT_{Y_4}, CPT_{Y_5} and CPT_S referred to by the specifications of CPTs corresponding to AND and OR relations

Table 7. Distribution parameters of basic units

No.	Parameters	No.	Parameters
X_1	β_1=β_2=2.769; η_1=η_2=[4794.4,5381.5]	X_6	μ_6=[7.2442,7.5700]; σ_6=0.1980
X_2		X_7	μ_7=[7.2442,7.5700]; σ_7=0.1980
X_3	β_3=6.02; η_3=[7439.4,7752.6]	X_8	μ_8=[8.4287,8.5937]; σ_8=0.1003
X_4	β_4=1.935; η_4=[8459.8,9746.6]	X_9	μ_9=[8.3428,8.4692]; σ_9=0.0768
X_5	β_5=8.33; η_5=[5851.9,5999.3]		

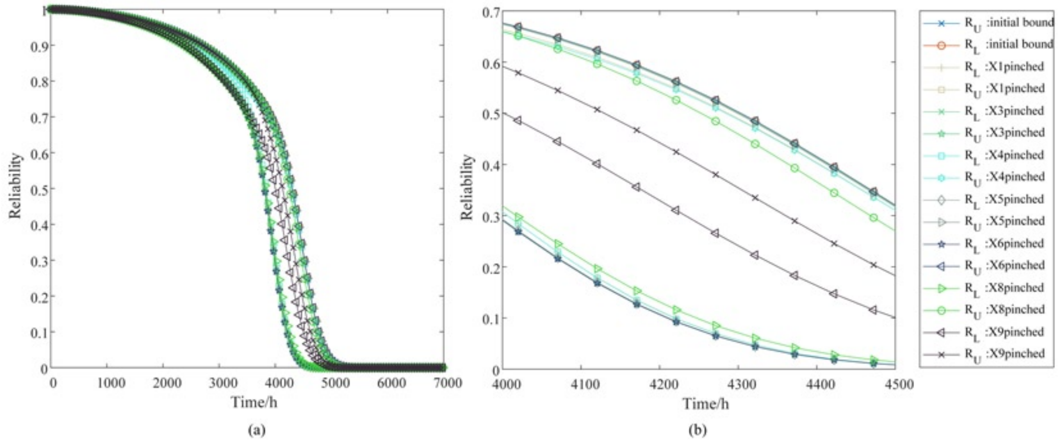


Fig. 14. P-boxes of different types of components with pinched uncertainty and initial bounds ((b) is the time interval (4000, 4500) of (a))

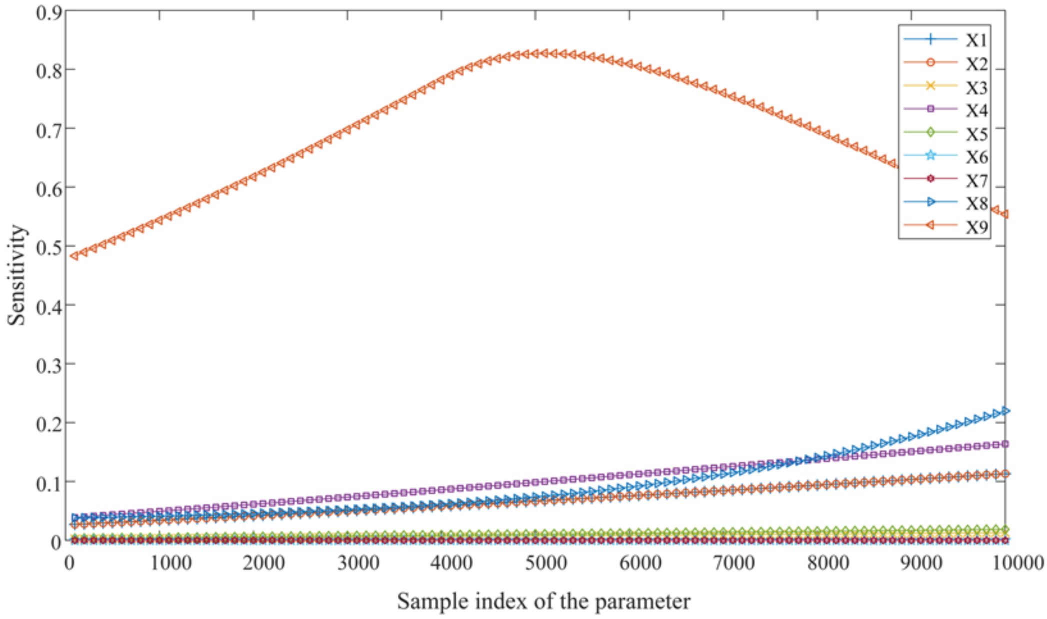


Fig. 15. Sensitivity deviation of each component

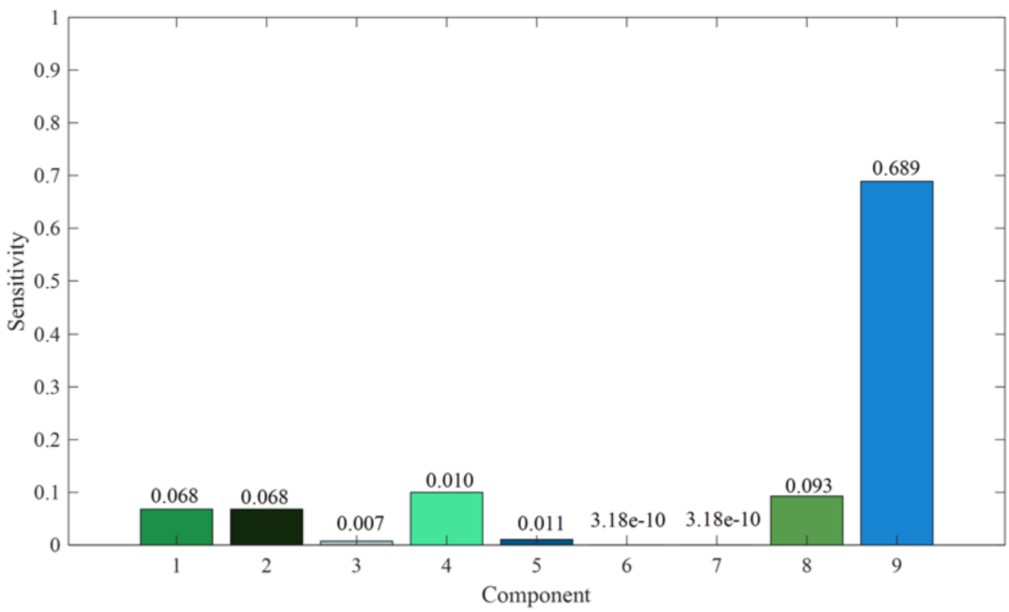


Fig. 16. Mean sensitivity of components $X_1 \sim X_9$

described in Section 2.1. Particularly, for the system, since X_6-X_7 and X_1-X_2 have the same probability distribution and are in the same connection, they can be regarded as equivalent according to Section 3. $f_{i,j}^k$ and $\overline{f_{i,j}^k}$ in the input vectors F and $F_{i,j}^k$ are obtained by the methods described in Table 7 and Section 3.2.1, where $l_k = 7$, $l_c = 9$ and $l_s = 10000$. The Bayesian forward reasoning and reliability characterization can be performed according to Eq. (11) and Eq. (12). On the other hand, due to the uncertainty in the input vector, the system reliability is uncertain and can be characterized by the p-box described in Section 3.3.

As shown in Fig. 14, there are several curves related to reliability bounds with each component pinched. It is obvious that component X_9 has the most critical impact on p-box uncertainty space compression after reducing the epistemic uncertainty of 9 types of components. However, it is difficult to distinguish the reduction effect of the other 8 components owing to the large number of curves crossing shown in Fig. 14. Hence, it is beneficial to further perform a sensitivity analysis of the system to resolve the aforementioned issues as well as obtain an accurate assessment indicator.

4.3. Sensitivity index and ranking

Based on Eq. (7) and the description of SA mentioned in Section 3.4, the epistemic uncertainty space size of the 10 sets of probability bounds in Fig. 14 should be estimated according to Eq. (6). Moreover, since it is difficult to determine the distribution of the target probability, the sensitivity might be biased. Therefore, to sample the parameter interval to offset the effect of the bias as described in Section 3.4, the sample size is chosen as l_s=10000. As depicted in Fig. 15, the ordinates, which refer to the sensitivity with the input vector of $F_{i,\mathrm{j}}^k$, are denoted as the reduction ratio of the uncertainty space with a maximum value of 1. From the results, it can be noted that the sensitivity changes of subsystem X_9, i.e., the hydraulic system, are nonlinear and much higher than those of X_1~X_8. Hence, the hydraulic system is the subsystem with the greatest uncertainty effect on the electromechanical system, which means that as the uncertainty of the system reliability must be reduced, a comprehensive analysis of the hydraulic system should be considered first. In contrast, the curves of X_3,X_5,X_6,X_7 are almost close to the x-axis with a large distance to X_1,X_2,X_4,X_8,X_9. Thus, it can be noted that X_3,X_5,X_6 and X_7 have the least impact on the system uncertainty, where the evaluation for system uncertainty reduction should be given the lowest consideration or even be deemed negligible. Furthermore, it is difficult to rank the sensitivity of X_1~X_8 since the curves are staggered with each other at similar amplitudes. Therefore, according to Eq. (8) and the method detailed in Section 3.4, we need to calculate the mean of each component sensitivity shown in Fig. 15 and plot the bar graph as shown in Fig. 16. Specifically, as seen in Fig. 16, the sensitivity of X_4, i.e., the reducer, is slightly higher than that of X_1, X_2 and X_8 but much higher than that of X_3, X_5, X_6 and X_7. Additionally, both X_3 and X_5 should be in the lower consideration of the system uncertainty reduction since their impact on the system uncertainty is just slightly higher than that of X_6 and X_7.

5. Conclusion

Sensitivity analysis has prominent application in the risk and reliability analysis field to explore how changes in the inputs of the component affect the outputs of the system. Nevertheless, the current SA study is mostly relevant to random variables with precise probability parameters, thus ignoring the existence of epistemic uncertainty. As industrial requirements increase, ISA theories have become popular solutions due to the inescapable imprecision in engineering. The target of the proposed method is to assess the reliability and sensitivity of mechatronic systems by considering the epistemic uncertainty and simultaneously accomplishing the sensitivity analysis.

In this paper, a pinching method was proposed to identify the sensitive components in a complex system on the basis of the reliability model established by the BN, and epistemic uncertainty is manipulated by the p-box. This method, on the basis of the BN reliability model, provides an alternative sensitivity index, unlike other methods such as Sobol indices [18], to successfully identify the components and subsystems with high sensitivity, which is an efficient way for engineers to reduce epistemic uncertainty. Moreover, compared with traditional Monte Carlo approaches, the brief concept and formulas of the p-box support a more intuitive SA and reduce the computational complexity. Two cases were applied to prove the feasibility, and we induce the sensitivity ranking via Fig. 10 and Fig. 16. Obviously, the accuracy of identifying the sensitive components is satisfactory. Particularly in the case of the electromechanical system from Ref. [10], the results show that the system epistemic uncertainty can be reduced by approximately 80% by pinching the uncertainty of the hydraulic system. Hence, efforts to reduce imprecision should primarily be made in hydraulic systems. During the analyses, this approach opens a new pathway based on the Bayesian network and pinching method in reliability sensitivity assessment, which indicates an efficient direction for mitigating engineering efforts in uncertainty reduction. In addition, during the analysis process, we also encountered several shortcomings. The sensitivity deviation cannot be totally eliminated by calculating the mean value. Therefore, our future work will focus on the selection and comparison of various sensitivity indices to improve the performance of the ISA method.

Acknowledgment

This work was partially supported by the National Natural Science Foundation of China under contract No. 51805073 and U1830207, the Chinese Universities Scientific Fund under contract No. ZYGX2018J061, the China Postdoctoral Science Foundation under contract No. 2015M582536, and the Sichuan Science and Technology Project under contract No. 2019JDJQ0015.

References

1. Guo H, Yang X. A simple reliability block diagram method for safety integrity verification. Reliability Engineering & System Safety 2007; 92(9): 1267-1273, https://doi.org/10.1016/j.ress.2006.08.002.
2. Simon C, Weber P, Evsukoff A. Bayesian networks inference algorithm to implement Dempster Shafer theory in reliability analysis. Reliability Engineering & System Safety 2008; 93(7): 950-963, https://doi.org/10.1016/j.ress.2007.03.012.
3. Simon C, Weber P, Levrat E. Bayesian Networks and Evidence Theory to Model Complex Systems Reliability. JCP 2007; 2(1): 33-43, https://doi.org/10.4304/jcp.2.1.33-43.
4. Ericson C A. Fault tree analysis. System Safety Conference, Orlando, Florida, 1999; 1: 1-9.
5. Pearl J. Fusion, propagation, and structuring in belief networks. Artificial Intelligence 1986; 29(3): 241-288, https://doi.org/10.1016/0004-

3702(86)90072-X.

6. Cai B, Liu Y, Liu Z, Tian X, Dong X, Yu S. Using Bayesian networks in reliability evaluation for subsea blowout preventer control system. Reliability Engineering & System Safety 2012; 108: 32-41, https://doi.org/10.1016/j.ress.2012.07.006.
7. Su C, Fu Y. Reliability assessment for wind turbines considering the influence of wind speed using bayesian network. Eksploatacja i Niezawodnosc -Maintenance and Reliability 2014; 16(1): 1-8.
8. Mi J, Li YF, Beer M, Broggi M, Cheng Y. Importance measure of probabilistic common cause failures under system hybrid uncertainty based on bayesian network. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(1): 112-120, https://doi.org/10.17531/ein.2020.1.13.
9. Bi S, Broggi M, Wei P, Beer M. The Bhattacharyya distance: Enriching the p-box in stochastic sensitivity analysis. Mechanical Systems and Signal Processing 2019; 129: 265-281, https://doi.org/10.1016/j.ymssp.2019.04.035.
10. Mi J, Li YF, Yang YJ, Peng W, Huang HZ. Reliability assessment of complex electromechanical systems under epistemic uncertainty. Reliability Engineering & System Safety 2016; 152: 1-15, https://doi.org/10.1016/j.ress.2016.02.003.
11. Xiahou T, Liu Y. Reliability bounds for multi-state systems by fusing multiple sources of imprecise information. IISE Transactions 2019; 0(0): 1-18, https://doi.org/10.1080/24725854.2019.1680910.
12. Dempster A P. The Dempster-Shafer calculus for statisticians. International Journal of approximate reasoning 2008; 48(2): 365-377, https://doi.org/10.1016/j.ijar.2007.03.004.
13. Misuri A, Khakzad N, Reniers G, Cozzani V. Tackling uncertainty in security assessment of critical infrastructures: Dempster-Shafer Theory vs. Credal Sets Theory. Safety Science 2018; 107: 62-76, https://doi.org/10.1016/j.ssci.2018.04.007.
14. Aguirre F, Sallak M, Schön W. Construction of Belief Functions From Statistical Data About Reliability Under Epistemic Uncertainty. IEEE Transactions on Reliability 2013; 62(3): 555-568, https://doi.org/10.1109/TR.2013.2273047.
15. Yang J, Huang HZ, He LP, Wen D, Zhu S-P. Failure Mode and Effects Analysis of Compressor Blades of Aeroengines Using Dempster-Shafer Evidence Theory. American Society of Mechanical Engineers Digital Collection: 2012: 863-870.
16. Mi J, Li YF, Peng W, Huang H-Z. Reliability analysis of complex multi-state system with common cause failure based on evidential networks. Reliability Engineering & System Safety 2018; 174: 71-81, https://doi.org/10.1016/j.ress.2018.02.021.
17. Feng G, Patelli E, Beer M, Coolen F P A. Imprecise system reliability and component importance based on survival signature. Reliability Engineering & System Safety 2016; 150: 116-125, https://doi.org/10.1016/j.ress.2016.01.019.
18. Schöbi R, Sudret B. Global sensitivity analysis in the context of imprecise probabilities (p-boxes) using sparse polynomial chaos expansions. Reliability Engineering & System Safety 2019; 187: 129-141, https://doi.org/10.1016/j.ress.2018.11.021.
19. Ferson S, Kreinovich V, Grinzburg L, S.Myers D. Constructing probability boxes and Dempster-Shafer structures. Sandia National Lab. (SNL-NM), Albuquerque, NM (United States) 2015.
20. Wei P, Lu Z, Song J. Variable importance analysis: A comprehensive review. Reliability Engineering & System Safety 2015; 142: 399-432, https://doi.org/10.1016/j.ress.2015.05.018.
21. Wei P, Lu Z, Ruan W, Song J. Regional sensitivity analysis using revised mean and variance ratio functions. Reliability Engineering & System Safety 2014; 121: 121-135, https://doi.org/10.1016/j.ress.2013.08.001.
22. Alvarez D A. Reduction of uncertainty using sensitivity analysis methods for infinite random sets of indexable type. International journal of approximate reasoning 2009; 50(5): 750-762, https://doi.org/10.1016/j.ijar.2009.02.002.
23. Sankararaman S, Mahadevan S. Separating the contributions of variability and parameter uncertainty in probability distributions. Reliability Engineering & System Safety 2013; 112: 187-199, https://doi.org/10.1016/j.ress.2012.11.024.
24. Krzykacz-Hausmann B. An approximate sensitivity analysis of results from complex computer models in the presence of epistemic and aleatory uncertainties. Reliability Engineering & System Safety 2006; 91(10): 1210-1218, https://doi.org/10.1016/j.ress.2005.11.019.
25. Oberguggenberger M, King J, Schmelzer B. Classical and imprecise probability methods for sensitivity analysis in engineering: A case study. International Journal of Approximate Reasoning 2009; 50(4): 680-693, https://doi.org/10.1016/j.ijar.2008.09.004.
26. Helton J C, Johnson J D, Oberkampf W L, Sallaberry C J. Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty. Reliability Engineering & System Safety 2006; 91(10): 1414-1434, https://doi.org/10.1016/j.ress.2005.11.055.
27. Ferson S, Troy Tucker W. Sensitivity analysis using probability bounding. Reliability Engineering & System Safety 2006; 91(10): 1435-1442, https://doi.org/10.1016/j.ress.2005.11.052.
28. Mi J, Cheng Y, Song Y, Bai L, Chen K. Application of dynamic evidential networks in reliability analysis of complex systems with epistemic uncertainty and multiple life distributions. Annals of Operations Research, 2019:1-23, https://doi.org/10.1007/s10479-019-03211-4.
29. Montgomery V. New statistical metliods in risk assessment by probability bounds 2009.
30. Huang X, Coolen F P A. Reliability sensitivity analysis of coherent systems based on survival signature. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 2018; 232(6): 627-634, https://doi.org/10.1177/1748006X18754974.
31. Feng G, George-Williams H, Patelli E, Coolen F P A, Beer M. An efficient reliability analysis on complex non-repairable systems with common-cause failures. Safety and Reliability-Safe Societies in a Changing World-Proceedings of the 28th International European Safety and Reliability Conference; 2018: 2531-2538, https://doi.org/10.1201/9781351174664-318.

He LIANG
School of Automation Engineering,
University of Electronic Science and Technology of China,
No.2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu, Sichuan, 611731, P.R. China

Jinhua MI
School of Automation Engineering,
Center for System Reliability and Safety,
University of Electronic Science and Technology of China,
No.2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu, Sichuan, 611731, P.R. China

Libing BAI
Yuhua CHENG
School of Automation Engineering,
University of Electronic Science and Technology of China,
No.2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu, Sichuan, 611731, P.R. China

Emails: harrisliang@outlook.com, jinhuami@uestc.edu.cn,
libing.bai@uestc.edu.cn, yhcheng@uestc.edu.cn

Article citation info:
MA B, YAN S, WANG X, CHEN J, ZHENG C. Similarity-based failure threshold determination for system residual life prediction. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 520–529, http://dx.doi.org/10.17531/ein.2020.3.15.

Biao MA
Shufa YAN
Xu WANG
Jianhua CHEN
Changsong ZHENG

SIMILARITY-BASED FAILURE THRESHOLD DETERMINATION FOR SYSTEM RESIDUAL LIFE PREDICTION

OKREŚLANIE PROGU AWARII NA PODSTAWIE PODOBIEŃSTWA JAKO METODA POZWALAJĄCA NA PRZEWIDYWANIE TRWAŁOŚCI RESZTKOWEJ SYSTEMU

An accurate determination of the system failure threshold is an essential requirement in achieving an appropriate system residual life prediction and a reasonable planned maintenance strategy optimization afterward for degradation systems. This paper proposes a failure threshold determination method based on quantitative measurement of the similarity between the operating system and the historical systems. The similarity is formulated by a weighted average function and then calculated by a convex quadratic formulation to minimizing the variance between the operating system and the historical systems. With an accurate determination of the system failure threshold in real-time, a better prediction of the residual life for the operating system is achieved. Finally, a real case study for several power-shift steering transmission systems monitored using oil spectral analysis is adopted to illustrate and numerically compare the improved performance of the proposed method.

Keywords: *system failure threshold; residual life; similarity; prognostics; oil field data.*

W przypadku systemów podlegających degradacji, dokładne określenie progu awarii systemu stanowi niezbędny warunek dokonania trafnej prognozy jego trwałości resztkowej oraz późniejszej optymalizacji strategii konserwacji rutynowych. W artykule zaproponowano metodę wyznaczania progu awarii opartą na ilościowym pomiarze podobieństwa między systemem użytkowanym obecnie a systemami użytkowanymi uprzednio. Podobieństwo formułuje się na podstawie funkcji średniej ważonej, a następnie oblicza na podstawie wypukłej formy kwadratowej w celu zminimalizowania wariancji między obecnie użytkowanym systemem a uprzednimi systemami. Dzięki dokładnemu określeniu progu awarii systemu w czasie rzeczywistym uzyskuje się lepszą prognostykę trwałości resztkowej obecnie użytkowanego systemu. W końcowej części pracy, w celu zilustrowania i numerycznego porównania ulepszonej wydajności proponowanej metody, zaprezentowano studium przypadku obejmujące kilka układów przeniesienia napędu monitorowanych przy użyciu analizy spektralnej oleju.

Słowa kluczowe: *próg awarii systemu; trwałość resztkowa; podobieństwo; prognostyka; dane z badań oleju.*

Abbreviations and acronyms

PSST	Power-shift steering transmission
PHM	Prognostics and health management
RL	Residual life
CDF	Cumulative distribution function
Mh	Motorhour

Nomenclature

$L_{j,t}$	Degradation data for system i at time t
φ_i	Random effects of the degradation model for system i
η	Functional form of the degradation model
$\varepsilon_{i,t}$	Random noise for system i at time t
$\omega_{i,j}$	Weight parameter corresponding to the historical system i and the operating system j
$\tilde{L}_{j,t}$	Reconstructed degradation model for system j at time t using the weighted average method
$\boldsymbol{L}_{j,\cdot}$	Collected real-time degradation data until time t
$\hat{\varphi}_i$	Estimated model parameters of historical systems
$\hat{L}_{j,t}$	Reconstructed degradation data for system j at time t using the Bayesian updating method
$\boldsymbol{\theta}_j^{(1)}$	Updated random effects of the degradation model for system j
D_j	Failure threshold for system j
L_{i,n_i}	Degradation data for system i at time n_i
n_j	Number of data collections of operating system j
λ	Tuning coefficient

$R(\cdot)$	Regularization function
$\boldsymbol{X}_t$	Time-dependent covariates at time t
p	Order of the polynomial model
$\boldsymbol{\mu}_j^0, \boldsymbol{\mu}_j^1$	Prior and posterior mean of random effects $\boldsymbol{\theta}_j$
Σ_j^0, Σ_j^1	Prior and posterior variance of random effects $\boldsymbol{\theta}_j$
ψ_j	Design matrix for operating system j
u_j^d	Mean of failure threshold for system j
v_j^d	Variance of failure threshold for system j
T_j	Actual RL for testing PSST system j
$\tilde{T}_j$	Estimated RL for testing PSST system j
$\Phi(\cdot)$	CDF of the standard normal function
RL_e	Mean prediction error of system RL
FT_e	Mean prediction error of system failure threshold
N	Number of testing PSST systems
D_j^{True}	Actual failure threshold for PSST system j

1. Introduction

Degradation-induced failure is an inevitable and natural phenomenon for various industrial devices and systems. After a certain extent of degradation, a system will run to failure such that it no longer functions, which may result in production downtime, severe economic loss, and safety problems afterward. For example, for power-shift steering transmission (PSST) systems are the most vital components in military vehicles, mining machines, and heavy industry, and they always degrade to failure frequently than other components [14, 15]. Their failures are critically hazardous and often lead to catastrophic consequences, and therefore, should be avoided. One useful approach to avoid unexpected failures is to conducting prognostics and health management (PHM), in which machine condition monitoring plays a foundational role and has aroused intensive concerns in academia and industry [27, 31]. With the collected degradation data (e.g., vibration signals, temperature information, and oil field data) from condition monitoring in real-time, the residual life (RL) of an operating system can then be estimated by conducting prognostic analysis.

In the area of PHM, it is commonly assumed that a system failure will occur once its degradation data cross the predetermined failure threshold [13, 17, 24]. Therefore, the RL can be estimated by comparing the degradation profile with the failure threshold. Accordingly, to accurately evaluate the RL distribution, two challenges must be addressed [21]: (1) A reasonable degradation model that can characterize the degradation profile of an operating system, so that the degradation mechanism can be accurately captured. (2) A reliable failure threshold that can reflect the failure mechanism of the system, so that the moment when the system no longer fulfills its functions can be accurately predicted. For many years, while numerous researches have been carried out focusing on the challenge (1) [2, 12, 26, 36], the existing research still insufficient to address the challenge (2), i.e., find a methodology that can precisely determine the failure threshold of the system. Therefore, the purpose of this paper is to address a failure threshold determination problem for the system with collected degradation data for the estimation of the system residual technical life.

Extensive work has been done in the application of different methods and techniques in degradation modeling, from which most of the current literature assumes that the collected degradation information can precisely characterize the underlying degradation mechanism of the system [1, 8]. In most cases, it is simply assumed that the failure threshold is known a priori, and, as a result, the failure threshold is always used as a fixed value for all systems. In practice, however, some recent research has shown that it may be irrational to use a fixed failure threshold [3, 4, 9]. For instance, as the work was done in [18, 22], the random failure threshold assumption has been adopted. Recently, the work in [29] presents a system failure threshold determination method based on the statistic characteristics of the last degradation data collections from multiple historical systems, and the uncertainty in the failure threshold distribution is also considered. However, the estimated failure threshold is obtained using an average of historical systems, which may not adequately consider the unique characteristics of each system. It is known that the degradation process of a system is stochastic, and is under the influence of many known and unknown factors. Therefore, the last collected degradation data always quite different, as shown in many research and applications [33, 37]. Besides, for some systems that are governed by multiple failure modes [19, 28], the failure threshold often different for each failure mode. As a result, the population-wide characteristic-based failure threshold determination method in [29] may be unable to fully consider the unique properties of an individual system. Therefore, a reliable failure threshold determination method that can extract the unique property in each individual system should be developed for the RL distribution estimation of a system when the failure threshold is not predetermined as a priori.

Motivated by the above observations, this paper proposes a failure threshold determination method based on quantitative measurement of the similarity between the operating system and historical systems. The similarity is formulated by using a weighted average function, and a convex quadratic formulation is then developed to minimizing the variance between the operating system and the historical systems. Unlike the existing PHM researches that assume the failure threshold as a deterministic value, in the proposed method, a random failure threshold is considered for different systems. Based on the proposed method, the failure threshold can be determined for an operating system with the collected real-time degradation information. This is of practical significance to attaining a more reasonable and accurate RL distribution estimation and, thus, is the main contribution of this paper. Finally, to illustrate the proposed method, a case study is provided for power-shift steering transmission systems.

The rest of this paper is organized as follows. Section II describes the details of the proposed method, which includes problem motivation, problem formulation, and the following residual life prediction procedures. Section III provides an illustrative real case study that involves a spectral oil data set from several power-shift steering transmission systems to show the effectiveness of the proposed method and its improved performance when used for system failure threshold determination and residual life prediction. Finally, Section IV draws the conclusion of this work and provides some future research directions.

2. Development of the methodology

This paper considers systems that degrade over time in working conditions, and condition monitoring techniques are conducted to collect the degradation data during the whole lifecycle (from an initial state to failure). Once the degradation data is collected, an associated degradation process $\{L(t),\ t \geq 0\}$ can be periodically analyzed to evaluate the degradation severity of the system. In engineering practice, many research efforts have been made on modeling the evaluation of the system degradation process and its relationship with the collected degradation data. For example, a polynomial degradation model introduced by Chinnam in [5], which has been commonly used in many applications [16, 23, 30, 32], can be written as follows:

$$L_{i,t} = \eta\left(\varphi_i, t\right) + \varepsilon_{i,t} \tag{1}$$

where $L_{i,t}$ represents the measurement of degradation data for system i at time t, η represents the functional form of the introduced degradation model, $\boldsymbol{\varphi}_i \in R^{p\times 1}$ represents the random effects of the degradation model for system i, and $\varepsilon_{i,t}$ is the random noise for system i at time t. With the introduced degradation model in Eq. (1), the degradation model of a system can be constructed based on the collected degradation data from condition monitoring during the lifecycle.

In the literature of PHM, the system degradation profiles (e.g., the failure modes, the failure threshold, and the constructed degradation model) of the same series of systems under the same working conditions (e.g., cycle condition, environment condition) are commonly assumed to be the same under some random variations [34, 35]. Under this assumption, our innovative idea is to reconstruct the operating system's degradation profile by the weighted average of the historical system degradation profiles. Specifically, the weight parameter measures the relative similarity between the degradation profiles of operating system j and historical system i compared to the other historical systems. Note that if the weight parameter $\boldsymbol{\omega}_{i,j}$ is obtained, the failure threshold of the operating system j can then be determined by the weighted sample moments of historical systems. By doing this, the system failure threshold can be online estimated by effectively using the condition monitoring data from a group of historical systems.

2.1. Description of the problem

According to the above mentioned innovative idea, a general description of the concerned system failure threshold determination problem is first provided in this section. Recall that the degradation model of an operating system can be reconstructed using the weighted average of the established historical system degradation models, and, if the weight parameter $\boldsymbol{\omega}_{i,j}$ is obtained, then the degradation model of an operating system can be reconstructed based on Eq. (1), which is given as:

$$E\left(\tilde{L}_{j,t} | \boldsymbol{L}_{j,\cdot}\right) = \sum_{i=1}^{m}\left(\omega_{i,j}\eta\left(\hat{\varphi}_i, t\right)\right) \tag{2}$$

where $\tilde{L}_{j,t}$ represents the reconstructed degradation model for operating system j at time t, and $L_{j,\cdot} = \left[L_{j,1},\ldots,L_{i,n_j}\right]^T$ represents the collected real-time degradation data during the lifecycle; $\omega_{i,j} \geq 0$ represents the proposed weight parameter corresponding to historical system i, and $\sum_{i=1}^{m}\omega_{i,j} = 1$; $\hat{\varphi}_i$ represents the estimated model parameters, m represents thenumber of historical systems.

On the other hand, when the real-time condition monitoring data are observed from an operating system, the Bayesian updating methods [7, 10] has also been widely used by many researchers to update the posterior distribution of the random effects for an operating system. In the perspective of the update process, the degradation model for operating system j can be updated using the Bayesian updating method based on the collected real-time condition monitoring data, which is given as:

$$E\left(\hat{L}_{j,t} | \boldsymbol{L}_{j,\cdot}\right) = \eta\left(E\left(\boldsymbol{\theta}_j^{(1)}\right), t\right) \tag{3}$$

where $\hat{L}_{j,t}$ represents the reconstructed degradation data for operating system j at time t using the Bayesian updating method, $\boldsymbol{L}_{j,\cdot}$ represents the collected real-time condition monitoring data until time t, and $\boldsymbol{\theta}_j^{(1)} \in R^{p\times 1}$ represents the estimated model parameters using the Bayesian updating method, which represents the updated random effects at time t.

Recall that the failure threshold of the same series of systems under the same operating conditions are expected to be the same, and the innovative idea is to reconstruct the degradation profile of an operating system by the weighted average of the historical system degradation profiles. In particular, if we consider the approximation of the Bayesian updated degradation model in Eq. (3) by using the reconstructed degradation model in Eq. (2), then, the system failure threshold distribution can be accurately estimated by using the information of historical systems with the real-time data of the operating system. In other words, the failure threshold for operating system j at the k th condition observation moment, D_j^k, can be determined by the weighted average of the last observation moments:

$$E\left[D_j^k\right] = \sum_{1}^{m}\left(\omega_{i,j}L_{i,n_i}^k\right) \tag{4}$$

where $\omega_{i,j} \geq 0$ for $i = 1,2,\ldots,m$, and $\sum_{i=1}^{m}\omega_{i,j} = 1$; L_{i,n_i} represents the degradation data of the last observation times before the failure of historical system i. Note that if the weight parameter $\omega_{i,j}$ is obtained, the failure threshold of operating system j can then be online determined. Therefore, the challenge here is to find the proper weight parameter, which will be specifically solved in the rest section.

2.2. Formulation of the methodology

Recall that the weight parameter $\boldsymbol{\omega}_{i,j}$ is proposed to measure the relative similarity between the degradation profile of operating system j and the degradation profile of historical system i compared to other historical systems. As a result, the aim of this section is to find the optimal weight parametersfor the operating system to maximize the goodness-of-fit between the reconstructed degradation model in Eq. (2) and the Bayesian updated degradation model in Eq. (3). In particular, the optimization model is formulated as a programming problem to minimize the sum of squared errors to estimate the optimal weight parameter $\boldsymbol{\omega}_{i,j}$, which is written as follows:

$$\begin{gathered}\min_{\omega_{i,j}} \sum_{t=1}^{n_j}\left(E\left[\hat{L}_{j,t}\right] - E\left[\tilde{L}_{j,t}\right]\right)^2 \\ \text{s.t. } \sum_{i=1}^{m}\omega_{i,j} = 1,\ \omega_{i,j} \geq 0,\ \text{for } i = 1,2,\ldots,m\end{gathered} \tag{5}$$

where n_j represents the number of data collections from operating system j, $E\left[\hat{L}_{j,t}\right]$ represents the degradation degree obtained by using the Bayesian updating method in Eq. (3), $E\left[\tilde{L}_{j,t}\right]$ represents the degradation degree obtained by using the weighted average method in Eq. (2) and $\sum_{i=1}^{m}\omega_{i,j}=1$ represents that the reconstructed degradation model is the weighted average of the historical system degradation profiles.

By solving the optimization problem in Eq. (5), the optimal weight parameter $\omega_{i,j}$ can be obtained and, the Bayesian updated degradation model can then be fitted by using the reconstructed degradation model of historical systems. It is worth note that overfitting problem may appear when the optimization model in Eq. (5) was directly implemented, especially when the operating system is in the initial stage that the number of available degradation data collections n_j is small. To solve this problem, a regularization function is introduced as:

$$R\left(\omega_{1,j},\ldots,\omega_{m,j}\right)=\sum_{i=1}^{m}\left(\omega_{i,j}\left\{E\left[\hat{L}_{j,n_i}\right]-L_{i,n_i}\right\}^2\right) \tag{6}$$

where $E\left[\hat{L}_{j,n_i}\right]$ represents the expected degradation data measurement for operating system j at the failure moment of historical system i, which can be obtained by using the Bayesian updating method in Eq. (3) up to the time n_i, L_{i,n_i} represents the measurement of degradation data at the failure moment of historical system i.

Remark 1: The proposed regularization function can be understood in this perspective: Recall that the main idea of our proposed failure threshold determination method is using the weighted average of the last measurements of degradation data of historical systems to determine the expected failure threshold of operating system j, as shown in Eq. (4). In this way, if the Bayesian updated degradation data for an operating system j at time n_i is obviously different from the measurement of degradation data for historical system i at time n_i, the historical system i has less impact on determining the failure threshold of operating system j. In other words, as shown in the regularization function of Eq. (6), we think a large penalty is supposed to be added to historical system i.

In summary, the concerned optimization model can be formulated by combining the regularization function in Eq. (6) with the formulation Eq. (5):

$$\begin{gathered}\min_{\omega_{i,j}}\sum_{t=1}^{n_j}\frac{\left(E\left[\hat{L}_{j,t}\right]-E\left[\tilde{L}_{j,t}\right]\right)^2}{n_j}+\lambda R\left(\omega_{1,j}\ldots,\omega_{m,j}\right)\\ \text{s.t. }\sum_{i=1}^{m}\omega_{i,j}=1,\ \omega_{i,j}\geq 0\text{ , for } i=1,2,\ldots,m\end{gathered} \tag{7}$$

where λ represents the tuning coefficient that measures the relative importance of the regularization function compared with the sum of squared errors between the reconstructed degradation model and the Bayesian updated degradation model. The tuning coefficient λ can be calculated with cross-validation [11], and in engineering practice, the value of λ often determined according to the importance assigned to each item.

Remark 2: Note that the optimization model in the Eq. (7) considers both the differences between the reconstructed degradation data for operating system j and the measurement of degradation data of historical systems at their failure moments, and the means squared error of the collected degradation data in the time domain for operating system j.

As noted above, the failure threshold for operating system j can be determined by using the optimal solution of weight parameters $\boldsymbol{\omega}_j$ with Eq. (4). Once the failure threshold D_j^k is determined, the RL for operating system j can then be evaluated. Therefore, in the next section, the RL distribution derivation method of an operating system will be investigated based on the proposed model in Eq. (7).

2.3. Estimation of the residual life

Recall that a system failure will occur once the degradation data cross the failure threshold. Without loss of generality, given the real-time degradation data $\boldsymbol{L}_{j,\cdot}=\left[L_{j,1},\ldots,L_{j,n_j}\right]^T$ of system j collected up to the current sampling moment n_j, the RL distribution of system j, $\tilde{T}_j$ can be estimated by:

$$P\left(\tilde{T}_j\leq \text{t}|\boldsymbol{L}_{j,\cdot}\right)=P(L_{j,n_j+t}\geq D_j\mid \boldsymbol{L}_{j,\cdot}) \tag{8}$$

where D_j represents the failure threshold of system j, L_{j,n_j+t} represents the measurement of project degradation data at the future sampling time n_j+t.

Therefore, to accurately evaluate the RL distribution of the operating system, a degradation model that can characterize the system degradation profile should be first established. In this paper, the polynomial function form of the degradation model is used for its useful mathematical properties [5]. To be specific, a widely used p th-order polynomial degradation model is given as:

$$L_{j,t}=\boldsymbol{X}_t\boldsymbol{\theta}_j+\varepsilon_{j,t}=\sum_{k=0}^{p}\theta_{j,k}t^k+\varepsilon_{j,t} \tag{9}$$

where $\boldsymbol{X}_t=\left[1,t,\ldots,t^p\right]$, and p represents the order of the degradation model; $\boldsymbol{\theta}_j$ is the vector that represents the random effects of the model that follows a multivariate normal distribution $N_{p+1}\left(\boldsymbol{\mu}_j,\Sigma_j\right)$, $\varepsilon_{j,t}$ represents the random noise and it is assumed to follow $N\left(0,\sigma_j^2\right)$. Note that after the degradation profiles of historical systems are fitted, then the prior distribution can be estimated:

$$\boldsymbol{\theta}_j^{(0)}\sim \boldsymbol{N}_{p+1}\left(\boldsymbol{\mu}_j^0,\Sigma_j^0\right), \tag{10}$$

where $\boldsymbol{\mu}_j^0$ is the prior mean value of the random effects, and Σ_j^0 is the prior variance value of the random effects. Then, the posterior distribution, $\boldsymbol{\theta}_j^{(1)}$, can be calculated by using the collected real-time degradation data $\boldsymbol{L}_{j,\cdot}$ of operating system j. Specifically, using the Bayesian updating approach that introduced in Makis *et al*. [36], the updated random effects $\boldsymbol{\theta}_j^{(1)}$ will be obtained as the following normal distribution:

$$\boldsymbol{\theta}_j^{(1)}=\boldsymbol{\theta}_j\mid \boldsymbol{L}_{j,\cdot}\sim N_{p+1}\left(\boldsymbol{\mu}_j^1,\Sigma_j^1\right) \tag{11}$$

where $\boldsymbol{\mu}_j^1=\left(\dfrac{\boldsymbol{\psi}_j^T\boldsymbol{\psi}_j}{\sigma_j^2}+\left(\Sigma_j^0\right)^{-1}\right)^{-1}\left(\dfrac{\boldsymbol{\psi}_j^T\boldsymbol{L}_{j,\cdot}}{\sigma_j^2}+\left(\Sigma_j^0\right)^{-1}\boldsymbol{\mu}_j^0\right)$,

$\Sigma_j^1=\left(\dfrac{\boldsymbol{\psi}_j^T\boldsymbol{\psi}_j}{\sigma_j^2}+\left(\Sigma_j^0\right)^{-1}\right)^{-1}$, and

$$\boldsymbol{\varnothing}_j \in R^{n_j \times (p+1)} = \begin{bmatrix} 1 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & t^p \\ \dots & \dots & \dots \\ 1 & \dots & n_j^p \end{bmatrix} \boldsymbol{\psi}_j \in R^{n_j \times (p+1)} = \begin{bmatrix} 1 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & t^p \\ \dots & \dots & \dots \\ 1 & \dots & n_j^p \end{bmatrix}.$$

Remark 3: Note that other distributions or simulation methodologies can also be used to calculate the posterior distribution and the RL of an operating system. Here, the use of the normal distribution for the model parameter characterizations is to take the convenience of its closed-form solution results. In fact, for many degradation models introduced in the existing literature, a normal distribution assumption for the random effects has been widely used [13, 17, 24, 31].

Remark 4: Note that other degradation models can also be utilized to build the degradation profiles and to determine the failure threshold. Here, the use of the polynomial form for the system degradation model is to take advantage of the convenient when calculating the parameters and the system RL. In fact, many other degradation models can also be transformed into the polynomial form [30,32], such as the random coefficient growth model [8], and the exponential form model [34].

Therefore, for the convenience of illustration, we only focus on the polynomial form for degradation models and the normal distribution for the random effects. Consequently, given the updated random effects $\boldsymbol{\theta}_j^{(1)}$, the degradation model of operating system j at time twill be obtained:

$$L_{j,t} \mid \boldsymbol{\theta}_j^{(1)} \sim N\left(\boldsymbol{X}_t \boldsymbol{\mu}_j^1, \boldsymbol{X}_t \Sigma_j^1 \boldsymbol{X}_t^T + \sigma_j^2\right) \tag{12}$$

Recall that the mean value u_j^d and variance value v_j^d of the failure threshold D_j for an operating system j can be calculated by $u_j^d = \sum_{i=1}^{m}\left(\omega_{i,j} L_{i,n_i}\right)$ and $v_j^d = \sum_{i=1}^{m}\left(\omega_{i,j} L_{i,n_i}^2\right) - \left(\sum_{i=1}^{m}\left(\omega_{i,j} L_{i,n_i}\right)\right)^2$ using Eq. (4). We further assume the system failure threshold D_j follows a normal distribution as well, i.e., $D_j \sim N\left(u_j^d, v_j^d\right)$. Note that this normal assumption for the system failure threshold distribution has been widely used in much existing research [13, 17, 24, 31]. Then, using the determined system failure threshold, the CDF of the RL $\tilde{T}_j$ for operating system j based on the collected real-time degradation data $\boldsymbol{L}_{j,\cdot}$ can be calculated as:

$$P\left(\tilde{T}_j \le t \mid \boldsymbol{L}_{j,\cdot}\right) = P\left(L_{j,n_j+t} \ge D_j \middle| \boldsymbol{L}_{j,\cdot}\right) = \Phi\left(\frac{\boldsymbol{X}_{n_j+t}\, \text{¼}_j^1 - \mu_j^d}{\sqrt{\boldsymbol{X}_{n_j+t} \Sigma_j^1 \boldsymbol{X}_{n_j+t}^T + \sigma_j^2 + v_j^d}}\right) = \Phi(\mathrm{g}(\mathrm{t})) \tag{13}$$

Given that the RL for operating system j should be greater than 0, the truncated CDF was further considered conditioning on $\tilde{T}_j \ge 0$:

$$P\left(\tilde{T}_j \le t \mid \tilde{T}_j \ge 0, \boldsymbol{L}_{j,\cdot}\right) = \frac{P\left(0 \le \tilde{T}_j \le t \middle| \boldsymbol{L}_{j,\cdot}\right)}{P\left(\tilde{T}_j \ge 0 \middle| \boldsymbol{L}_{j,\cdot}\right)} = \frac{\Phi(\mathrm{g}(\mathrm{t})) - \Phi(\mathrm{g}(0))}{1 - \Phi(\mathrm{g}(0))} \tag{14}$$

Since the truncated CDF in Eq. (14) is skewed, the median can be used as the point estimator for the RL prediction.

2.4. Flowchart of the methodology

The flowchart of the proposed system failure threshold determination method is shown in Fig. 1. A convex quadratic formulation is developed to combine the information from the degradation data of historical systems and the information from the real-time condition monitoring data of an operating system. Using the degradation data from historical systems, the degradation models of each historical system is first established and then reconstructed using a weighted average method to fit the degradation model of an operating system, as shown in Eq. (2). In addition, the information about the last condition monitoring data collections ($L_{i,t}$) is extracted, and in other words, the failure threshold of historical system i is estimated using the last condition monitoring data collection ($L_{i,t}$). Then, to established the real-time update model, the prior distribution of the parameters in the Bayesian updated model is also obtained using the historical systems. Then, using the degradation data from the operating system, the degradation model for the operating system is updated based on the Bayesian approach using the collected real-time condition monitoring data, as shown in Eq. (11). Particularly, the method focuses on using the degradation profiles of historical systems to fit the degradation profile of an operating system. And then, using the updated degradation model and the reconstructed degradation model of historical systems, the failure threshold of the operating system is finally determined by solving Eq. (7). Finally, using the determined failure threshold and the

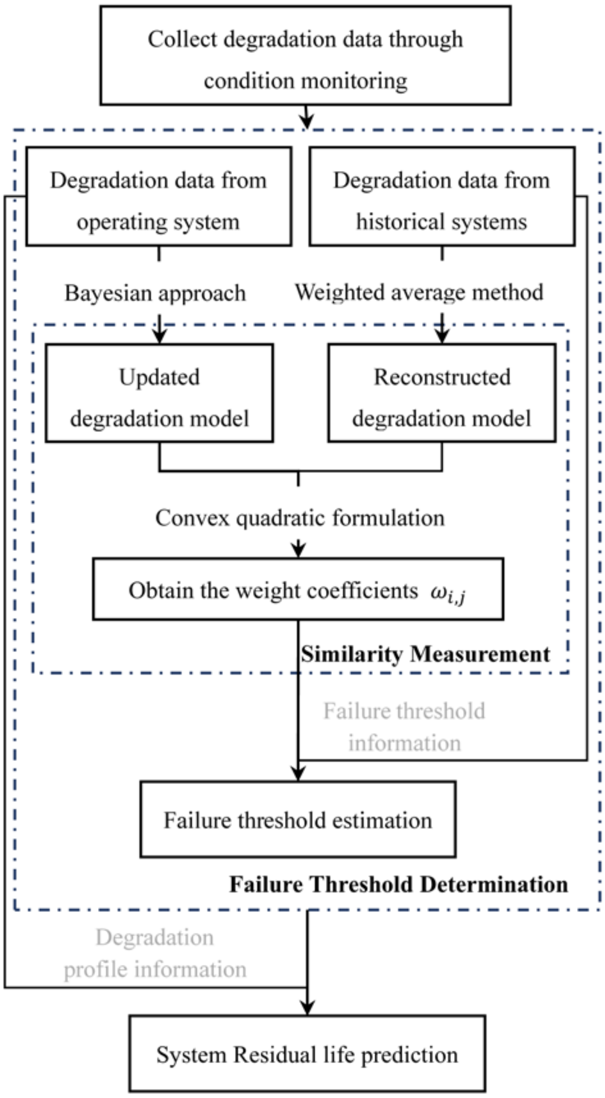


Fig. 1. Failure threshold determination framework for RL prediction

collected real-time condition monitoring data, the degradation model of an operating system can be established, and the system RL can then be estimated by Eq. (13) and Eq. (14).

3. Imperfect PM model considering system aging

In this section, the performance of the proposed system failure threshold determination method is investigated using a degradation dataset from several PSST systems. Specifically, we compare the prognostic performance of the proposed method with the existing population-wide characteristic-basedsystem failure threshold determination method. In fact, the method in [29] can be regarded as a special case of the proposed method in which each historical system is treated with equal importance, i.e., the weight coefficient is set as $\omega_{i,j}=1/\mathrm{m}$, for $i=1,2,\ldots,\mathrm{m}$. On the contrary, when determining the failure threshold of the operating system j, the proposed method will calculate the optimal weight coefficient $\omega_{i,j}$ for $i=1,2,\ldots,\mathrm{m}$ by solving Eq. (7).

3.1. Origin of the data

The dataset is collected from several PSST systems (Fig.2), which is a widely used mechanical powertrain system in heavy tracked vehicles. With PSST operating, metal debris mixed in lubricating oil accelerates the wear of every mechanical component, consequently leads to the degradation of the PSST system. However, the underlying degradation is not directly observable and can be only indirectly assessed via oil spectral analysis, which is a commonly used technique to monitor the underlying degradation conditions in oil-lubricated machines [24, 25, 28, 32]. In addition, the underlying failure mechanism in PSSTis not explicitly revealed. Therefore, researchers have to rely solely on the spectral oil data collected in the oil field dataset to establish the PSST degradation model and predict the RL.

In particular, the oil field dataset obtained from the reliability testbed (Fig. 3) for the PSST system is used for the case study illustration. Each system was run to failure under the same operating condition: cyclic multi-gear, load-varying, and multispeed. After a period of more than 10 years, we have collected oil field data for more than one thousand samples, and a detailed description of the dataset can be found in [31, 32]. The dataset we used in this paper contains $m=45$ training systems and 15 testing systems, of which each system contains more than 30 oil samples that were collected from 0 **Mh** to up to 284**Mh** according to the sample periodof nearly 5 **Mh**. Each sample contains 15 types of spectral oil data, among which 6 types, namely, Cr, Ni, Cu,

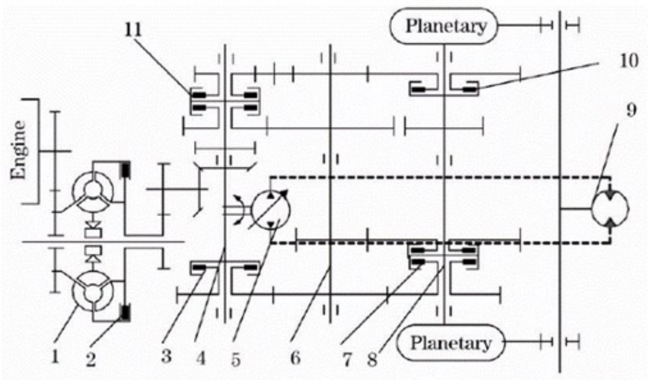


Fig. 2. Sketch of the PSST
1: Hydraulic torque convertor. 2: CV clutch.3: CH clutch. 4: First-shaft. 5: Steering pump. 6: Second shaft. 7: C1C2 clutch.8: Thirdshaft. 9: Steering motor. 10: C3 clutch. 11: CLCR clutch

Mn, Fe, and Mo, are selected in [30] and shown to be highly related to the degradation mechanism. Due to space restrictions, the selected spectral oil data of one PSST system are shown in Fig. 4.

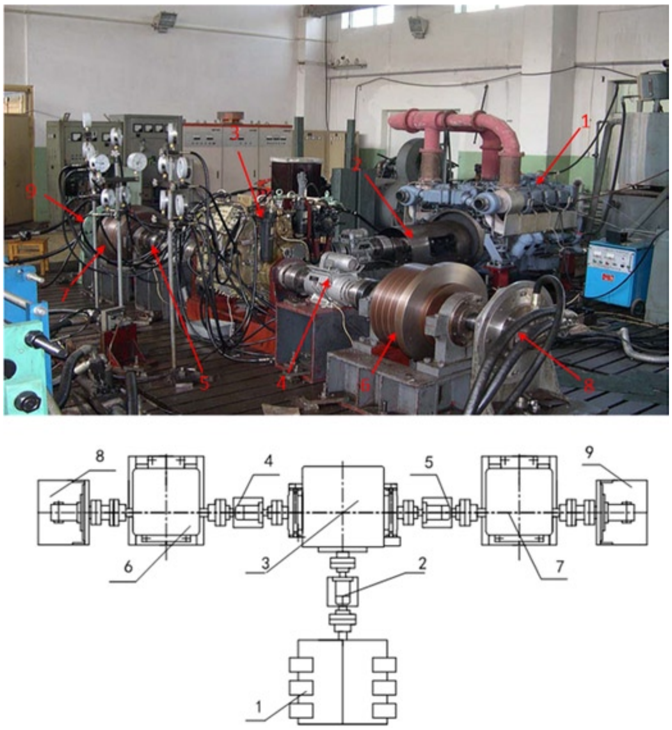


Fig. 3. A life-cycle testbed of the PSST
1: Diesel engine. 2, 4, 5: Torque and speed sensors.3: PSST. 6, 7: Inertia discs. 8, 9: Loading piston pump

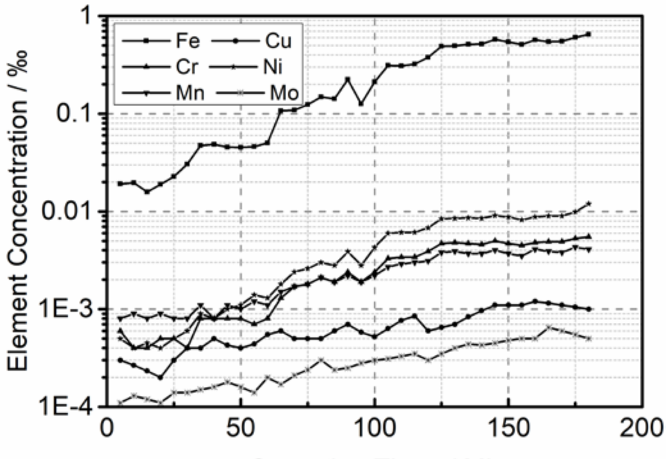


Fig. 4. Spectral oil data for one PSST system

3.2. System degradation modeling

For each PSST system, there are 6 types of spectral oil data that monitor the degradation process of the PSST system performance. Using these spectral oil data, the system degradation model can be established. The concerned PSST systems are subjected to three potential failure modes: 1) a fault at the transmission gears, or 2) a fault at the wet clutches, or 3) a fault at the rotary sealings. According to the previous research in [20], the spectral oil data following an exponential functional form. Thus, the following exponential degradation model is considered to model the degradation profiles of PSST systems:

$$y_{i,j,t}=\gamma_j+e^{\theta_{i,j,0}+\theta_{i,j,1}t+\theta_{i,j,2}t^2+\epsilon_{i,j,t}} \tag{15}$$

where $y_{i,j,t}$ represents the degradation data measurement for PSST system i, spectral oil data j at time t; γ_j represents the initial-effect coefficient for spectral oil data j, $\epsilon_{i,j,t}$ represents the random noise; and $\boldsymbol{\theta}_{i,j,0}$, $\boldsymbol{\theta}_{i,j,1}$ and $\boldsymbol{\theta}_{i,j,2}$represent the random effects for PSST system i, spectral oil data j .Similar to Makis *et al*. [6] and Liu *et al*. [20], a log-transformation is used to process the original spectral oil data, and then the logged spectral oil data is modeled as follows:

$$L_{i,j,t} = \ln\left(y_{i,j,t} - \gamma_j\right) = \theta_{i,j,0} + \theta_{i,j,1}t + \theta_{i,j,2}t^2 + \epsilon_{i,j,t} \tag{16}$$

Using the established degradation model in Eq. (16), the PSST system RL can be then determined. It is noted that all the selected 60 PSST systems are degraded and failed according to one of the three possible failure modes. In addition, it is assumed that all PSST systems within the same failure modes following the assumptions: 1) have the same expected failure threshold subject to the same variations; and 2) the random effects have the same distribution.

To challenge our proposed method, it is further assumed that the failure modes of the training PSST system and the operating PSST system are unknown, which is usually the case in the real application. This assumption is to show the superiority of our proposed method in inferring the failure threshold and the RL of an operating PSST system, even if the failure mode of the training PSST system is unknown. In this way, the performance of our proposed system failure determination method can be evaluated and compared with the existing method [29]. Specifically, two metrics listed in the following are used for evaluation:

1) The mean prediction error of the system RL (RL_e), which is defined as follows:

$$RL_e(\%) = \frac{100}{N} * \sum_{j=1}^{N} \frac{\left|\left(T_j + n_j\right) - \left(\tilde{T}_j + n_j\right)\right|}{T_j + n_j} = \frac{100}{N} * \sum_{j=1}^{N} \frac{T_j - \tilde{T}_j}{T_j + n_j} \tag{17}$$

where N represents the number of testing PSST systems, n_j represents the number of collected spectral oil data for testing PSST system j, and $T_j + n_j$ represents the lifetime for testing PSST system j; T_j represents the actual RL for testing PSST system j ,and $\tilde{T}_j$ represents the estimated system RL for the same testing PSST system.

2) The mean prediction error of the system failure threshold (FT_e), which is defined as follows:

$$FT_e(\%) = \frac{100}{N} * \sum_{j=1}^{N} \frac{\left|E\left[D_j\right] - D_j^{\mathrm{True}}\right|}{D_j^{\mathrm{True}}} \tag{18}$$

where $E\left[D_j\right]$ represents the estimated system failure threshold for testing PSST j using our proposed method in Eq. (4), and D_j^{True} represents the actual failure threshold for the same testing PSST system.

Matric (1) measures the mean prediction error between the predicted system residual life and the actual system residual life for the testing PSST system, and matric (2) measures the mean prediction error between the estimation value and the corresponding actual value of the system failure threshold for the testing PSST system. Note that these two metrics measure the concerned two aspects, namely, the failure threshold and system RL [4], when using collected condition monitoring data for system RL prediction.

3.3. System Failure Threshold Determination

For the system failure threshold determination, the procedures described in Section 2.4 are adopted. First, the system failure threshold of training PSST system i is estimated by the last observation (s_{i,n_i}) of the spectral oil data. Then the degradation model parameters are denoted as $\varphi_i = \left[\varphi_{i,0}, \varphi_{i,1}, \varphi_{i,2}\right]^T$ and the degradation model for each training PSST system i is reconstructed based on a second-order polynomial model. Next, the model parameters of testing system j are assumed to follow a normal distribution $\theta_j^{(0)} \sim N_{p+1}\left(\boldsymbol{u}_j^0, \Sigma_j^0\right)$, where $\boldsymbol{u}_j^0$ is estimated by the mean value of the parameters φ_i of training PSST systems, and Σ_j^0 is estimated by the variance value of the parameters φ_i of training PSST systems. Then the posterior distribution $\boldsymbol{\theta}_j^1$ is calculated using the real-time collected spectral oil data $\boldsymbol{L}_{j,\cdot}$ from the testing PSST system as shown in Eq. (11). Finally, the weight coefficient $\omega_{i,j}$ is calculated by solving Eq. (7) with five-fold cross-validation and consequently, the failure threshold of testing PSST system j can be estimated with the mean value $u_j^d = \sum_{i=1}^{m}\left(\omega_{i,j}s_{i,n_i}\right)$ and variance value $u_j^d = \sum_{i=1}^{m}\left(\omega_{i,j}s_{i,n_i}^2\right) - \left(\sum_{i=1}^{m}\left(\omega_{i,j}s_{i,n_i}\right)\right)^2$. Please note that Henze-Zircller's test is conducted in this case study based on the spectral oil data sets from the training PSST systems, and the calculation results show that using the normal distribution as the prior distribution for the random effects is satisfactory, as mentioned in Eq. (10).

The average prediction error of the failure threshold of the PSST system, FT_e as defined in Eq. (18), is shown in Fig.5, of which the best performance spectral oil data sample (Fe) is used for illustration. The threshold is modeled as a function of the operating time of the testing PSST systems. To be specific, the point corresponding to the "0" label is the average prediction error of the testing PSST systems in the initial state, while the point corresponding to the "50" label is the average prediction error when the testing PSST systems operating to 50% of the whole life. Here, for the convenience of illustration, the existing population-wide characteristic-based system failure threshold determination method in [29] is denoted as the "benchmark method".

From Fig.5, we can observe that our proposed method has an excellent performance, and the average prediction error is consistently reduced with the increase of the operating time. On the contrary, the benchmark method has poor performance, and the RL prediction accuracy does not improve with the increase of the operating time. It can be concluded that the proposed method provides a better result compared with the benchmark method when used for system failure

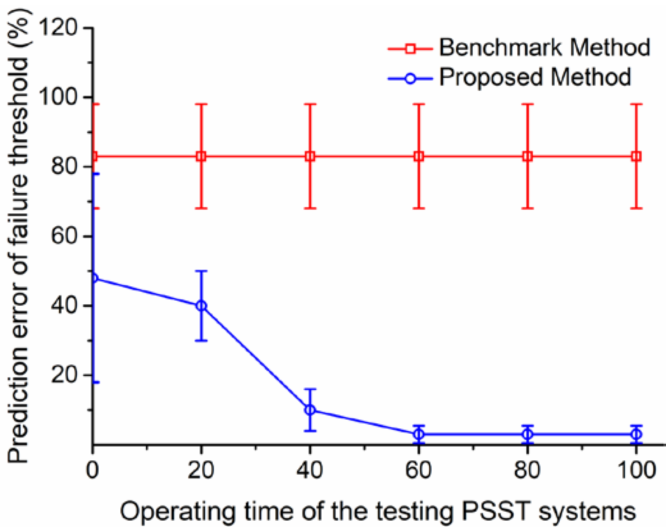


Fig. 5. Prediction error of the PSST system failure threshold using Fe

threshold determination. That is because that the threshold determination using the benchmark method just considering the population-wide characteristics from historical PSST systems, does not consider the individual characters from the operating PSST system. In addition, when using our proposed method, the average prediction error is approximately less than 3% when the operating PSST systems operate to half-life. To be specific, the result shows that the proposed method can determine the PSST system failure threshold accurately even the PSST system is operating in a stable condition. Moreover, from Fig.5, it can be seen that the average prediction error tends to be large when the testing PSST systems operate at an early stage (less than 20% of the lifetime). One possible reason is that the system degradation model in Eq. (3) is established using the Bayesian update method that may not fully capture the unique degradation features of the testing PSST systems when there are fewer spectral oil data samples available. However, when more spectral oil data samples are available, the accuracy of the PSST system failure threshold determination is improved significantly, which shows the rationality and superiority of the proposed system failure threshold determination method.

3.4. System residual life estimation

To further evaluate the performance of the proposed method for system RL prediction, the RL of the testing PSST systems is estimated by adopting the calculated system failure threshold in Section 3.3 and the proposed method in Section 2.3. Given the spectral oil data from all of the 15 testing PSST system, the average prediction error of the RL, as defined in Eq. (17), is then calculated by using the actual RL of the testing PSST system. The calculation result is summarized in Tab. 1.

From Tab. 1, we can observe that the accuracy of RL prediction using the proposed method is improved for all selected spectral oil data when compared with the benchmark method. Moreover, the average prediction error at different levels of actual RL for the testing PSST systems are calculated and shown in Fig.6, and the comparison between the proposed method and the benchmark method is shown in Fig.7. In Fig.6 and 7, the label “30” represents the average prediction error for the testing PSST system that has 30 or less actual RL. The minimum actual RL among all the testing PSST systems is 150. The label “150” represents the average prediction error using all the testing PSST systems. Thus, the same sets of testing PSST systems are used to calculate the accuracy of RL prediction as different levels of the actual RL.

From Fig.6, it can be seen that the RL prediction becomes accurate in general with the testing PSST system operating from the initial state to failure. This may because more spectral oil data samples are collected when the actual RL becomes less, and thus, the estimated system failure threshold, as well as the updated Bayesian degradation model for the operating PSST system, becomes more confidential.

Fig. 7 further shows the comparison by using the best performance spectral oil data sample (Fe) and the worst performance spectral oil data sample (Mn) according to the result in Tab.1. Based on Fig. 7, it can be clearly seen that the proposed system failure threshold determination method outperforms the original benchmark method in both spectral oil data. This observation further shows the rationality and superiority of our proposed method for system RL estimation.

Table 1. The average prediction error of the RL

Spectral oil data	Cr	Ni	Cu	Mn	Fe	Mo	Average
The proposed method	9.75%	10.6%	8.32%	15.85%	7.5%	12.6%	10.77%
The benchmark method	15.78%	14.55%	14.52%	20.21%	13.8%	18.06%	16.15%
Improvement	6.03%	3.95%	6.2%	4.36%	6.3%	5.46%	5.38%

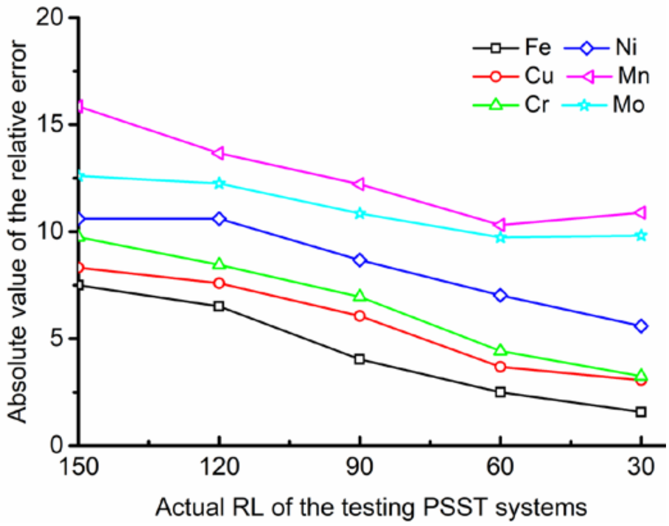


Fig. 6. Prediction error for the testing PSST systems using the proposed method

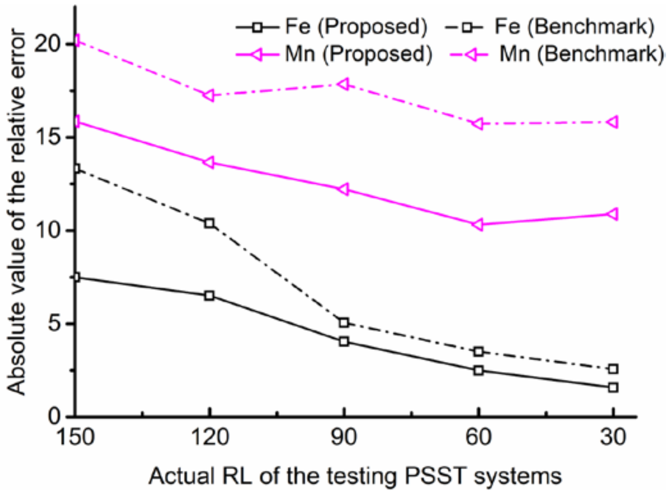


Fig. 7. Prediction error between the proposed method and the benchmark method

By comparing Fig. 6 and 7, it can be seen that the accuracy of RL prediction is highly related to the accuracy of system failure threshold determination. To be specific, with the improvement of the accuracy of system failure threshold determination for the operating PSST system, the accuracy of RL prediction using the proposed method also improves compared with the benchmark method.

4. Conclusion and discussion

This paper proposes a similarity-based failure threshold determination method for the system residual life prediction. To be specific, the similarity between the operating system and historical systems is measured by a weighted average function and calculated by using a convex quadratic formulation, and thus, the corresponding system failure threshold can be determined. The novelty of the proposed method is that the degradation profiles of historical systems and the real-time condition monitoring data of an operating system are combined for real-time estimating the system failure threshold of an operating system when the failure threshold is unknown.

The validity of the proposed method was demonstrated using selected oil field data collected from several PSST systems. The results not only show the effectiveness of the proposed method but also reveal the importance of real-time estimating the failure threshold for an operating system.

Compared with the existing failure threshold determination method which using population-wide characteristics, the proposed method provides an accurate determination of the system failure threshold, as illustrated in the case study. With the improved determination of the failure threshold, a more accurate system residual life estimation can then be obtained; even the monitored system is at an early moment of the degradation. Such advantages can be effectively used to obtain a better prognostic application and the following planned maintenance optimization. Following the results of the proposed method, the case study described in the existing works of literature, for example, Vališ et al. [24], Yan et al. [32] and Liu et al. [20], might be complemented when these methods are utilized for system soft failure prediction, health index extraction and other condition monitoring applications.

The main contribution of this paper is not only a new direction in the failure threshold determination for degradation systems but also open up possibilities for combining the real-time degradation data with the historical condition monitoring data. There are several possible directions for future research. First, other methods, such as a Bayesian framework, might be considered to update the failure threshold, as compared to the proposed frequentist method. Second, other degradation modeling method can fuse multiple condition monitoring data may have to be used when modeling other systems.

Acknowledgment

This work is partiallysupported by the NSFC under grant numbers51475044, 51975047, and partially supported by the China Scholarship Council under grant number 201806030083.The authors have no conflicts of interest to declare.

References

1. Alaswad S, Xiang Y. A review on condition-based maintenance optimization models for stochastically deteriorating system. Reliability Engineering & System Safety 2017; 157: 54-63, https://doi.org/10.1016/j.ress.2016.08.009.
2. Bian L, Gebraeel N, Kharoufeh J P. Degradation modeling for real-time estimation of residual lifetimes in dynamic environments. IIE Transactions 2015; 47(5): 471-486, https://doi.org/10.1080/0740817X.2014.955153.
3. Caballé N C, Castro I T, Pérez C J, Lanza-Gutiérrez J. M. A condition-based maintenance of a dependent degradation-threshold-shock model in a system with multiple degradation processes. Reliability Engineering & System Safety 2015; 134: 98-109, https://doi.org/10.1016/j.ress.2014.09.024.
4. Chehade A, Bonk S, Liu K. Sensory-based failure threshold estimation for remaining useful life prediction. IEEE Transactions on Reliability2017; 66(3): 939-949, https://doi.org/10.1109/TR.2017.2695119.
5. Chinnam R B. On-line reliability estimation for individual components using statistical degradation signal models. Quality and Reliability Engineering International 2002; 18(1): 53-73, https://doi.org/10.1002/qre.453.
6. Du Y, Wu T, Makis V. Parameter estimation and remaining useful life prediction of lubricating oil with HMM. Wear 2017; 376: 1227-1233, https://doi.org/10.1016/j.wear.2016.11.047.
7. Elwany A H, Gebraeel N Z. Sensor-driven prognostic models for equipment replacement and spare parts inventory. IIE Transactions 2008; 40(7): 629-639, https://doi.org/10.1080/07408170701730818.
8. Giraitis L, Kapetanios G, Yates T. Inference on multivariate heteroscedastic time varying random coefficient models. Journal of Time Series Analysis 2018; 39(2): 129-149. https://doi.org/10.1111/jtsa.12271.
9. Keizer M C O, Flapper S D P, Teunter R H. Condition-based maintenance policies for systems with multiple dependent components: A review. European Journal of Operational Research 2017; 261(2): 405-420, https://doi.org/10.1016/j.ejor.2017.02.044.
10. Kim M J, Jiang R, Makis V, Lee C G. Optimal Bayesian fault prediction scheme for a partially observable system subject to random failure. European Journal of Operational Research 2011; 214(2): 331-339, https://doi.org/10.1016/j.ejor.2011.04.023.
11. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In IJCAI 1995; 14(2): 1137-1145.
12. 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, https://doi.org/10.17531/ein.2019.4.18.
13. Lee J, Wu F, Zhao W, Ghaffari M, Liao L, Siegel D. Prognostics and health management design for rotary machinery systems-Reviews, methodology and applications. Mechanical systems and signal processing 2014; 42(1-2): 314-334, https://doi.org/10.1016/j.ymssp.2013.06.004.
14. Lei Y, Li N, Guo L, Li N, Yan T, Lin J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. Mechanical Systems and Signal Processing 2018;104: 799-834, https://doi.org/10.1016/j.ymssp.2017.11.016.
15. Li X, Makis V, Zuo H, Cai J. Optimal Bayesian control policy for gear shaft fault detection using hidden semi-Markov model. Computers & Industrial Engineering 2018; 119: 21-35, https://doi.org/10.1016/j.cie.2018.03.026.
16. Liao L. Discovering prognostic features using genetic programming in remaining useful life prediction. IEEE Transactions on Industrial Electronics 2013; 61(5): 2464-2472, https://doi.org/10.1109/TIE.2013.2270212.
17. Liu, K., Gebraeel, N. Z., Shi, J. A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis. IEEE Transactions on Automation Science and Engineering 2013; 10(3): 652-664. https://doi.org/10.1109/TASE.2013.2250282
18. Liu K, Huang S. Integration of data fusion methodology and degradation modeling process to improve prognostics. IEEE Transactions on Automation Science and Engineering 2014;13(1): 344-354, https://doi.org/10.1109/TASE.2014.2349733.
19. Liu, X., Li, J., Al-Khalifa, K. N., Hamouda, A. S., Coit, D. W., Elsayed, E. A. Condition-based maintenance for continuously monitored degrading systems with multiple failure modes. IIE transactions 2013; 45(4): 422-435. https://doi.org/10.1080/0740817X.2012.690930
20. Liu Y, Ma B, Zheng C S, Xie S Y. Failure prediction of power-shift steering transmission based on oil spectral analysis with Wiener process. Spectroscopy and Spectral Analysis 2015; 35(9): 2620-2624.
21. Okoh C, Roy R, Mehnen J, Redding L. Overview of remaining useful life prediction techniques in through-life engineering services. Procedia CIRP 2014; 16: 158-163, https://doi.org/10.1016/j.procir.2014.02.006.
22. Tang S J, Yu C Q, Feng Y B, Xie J, Gao Q H, Si X S. Remaining useful life estimation based on Wiener degradation processes with random failure threshold. Journal of Central South University 2016; 23(9): 2230-2241, https://doi.org/10.1007/s11771-016-3281-z.

23. Tian Z, Wong L, Safaei N. A neural network approach for remaining useful life prediction utilizing both failure and suspension histories. Mechanical Systems and Signal Processing 2010; 24(5): 1542-1555, https://doi.org/10.1016/j.ymssp.2009.11.005.
24. Vališ D, Žák L, Pokora O, Lánský P. Perspective analysis outcomes of selected tribodiagnostic data used as input for condition based maintenance. Reliability Engineering & System Safety 2016; 145: 231-242, https://doi.org/10.1016/j.ress.2015.07.026.
25. Wakiru J M, Pintelon L, Muchiri P N, Chemweno P K. A review on lubricant condition monitoring information analysis for maintenance decision support. Mechanical Systems and Signal Processing 2019; 118: 108-132, https://doi.org/10.1016/j.ymssp.2018.08.039.
26. Wang D, Tsui K L. Brownian motion with adaptive drift for remaining useful life prediction: Revisited. Mechanical Systems and Signal Processing 2018; 99: 691-701, https://doi.org/10.1016/j.ymssp.2017.07.015.
27. Wang J, Makis V, Zhao X. Optimal condition-based and age-based opportunistic maintenance policy for a two-unit series system. Computers & Industrial Engineering2019; 134: 1-10, https://doi.org/10.1016/j.cie.2019.05.020.
28. Xiao N, Huang H Z, Li Y, He L, Jin T. Multiple failure modes analysis and weighted risk priority number evaluation in FMEA. Engineering Failure Analysis 2011; 18(4): 1162-1170, https://doi.org/10.1016/j.engfailanal.2011.02.004.
29. Yan S, Ma B, Zheng C. Degradation index construction methodology for mechanical transmission based on fusion of multispectral oil data. Industrial Lubrication and Tribology 2019; 71(2): 278-283, https://doi.org/10.1108/ILT-04-2018-0154.
30. Yan S, Ma B, Zheng C. Health index extracting methodology for degradation modelling and prognosis of mechanical transmissions. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 137-144, https://doi.org/10.17531/ein.2019.1.15.
31. Yan S, Ma B, Wang X, Zheng C. Maintenance policy for oil-lubricated systems with oil analysis data. Eksploatacja i Niezawodnosc -Maintenance and Reliability 2020; 22(3): 455-464, https://doi.org/10.17531/ein.2020.3.8.
32. Yan S F, Ma B, Zheng C S. Remaining useful life prediction for power-shift steering transmission based on fusion of multiple oil spectra. Advances in Mechanical Engineering 2018; 10(6): 1687814018784201, https://doi.org/10.1177/1687814018784201.
33. Yan S F, Ma B, Zheng C S, Chen M. Weighted evidential fusion method for fault diagnosis of mechanical transmission based on oil analysis data. International Journal of Automotive Technology 2019; 20(5): 989-996, https://doi.org/10.1007/s12239-019-0093-9.
34. Yan S F, Ma B, Zheng C S, Zhu L A, Chen J W, Li H Z. Remaining useful life prediction of power-shift steering transmission based on uncertain oil spectral data. Spectroscopy and Spectral Analysis 2019; 39(2): 553-558.
35. Ye Z S, Xie M. Stochastic modelling and analysis of degradation for highly reliable products. Applied Stochastic Models in Business and Industry 2015; 31(1) 16-32, https://doi.org/10.1002/asmb.2063.
36. Zhai Q, Ye Z S. Degradation in common dynamic environments. Technometrics 2018; 60(4): 461-471, https://doi.org/10.1080/00401706.2017.1375994.
37. Zhang C, Lim P, Qin A K, Tan K C. Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics. IEEE transactions on neural networks and learning systems 2016; 28(10): 2306-2318, https://doi.org/10.1109/TNNLS.2016.2582798.
38. Zhang Z, Si X, Hu C, Lei Y. Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods. European Journal of Operational Research2018; 271(3): 775-796, https://doi.org/10.1016/j.ejor.2018.02.033

Biao MA
Shufa YAN
Xu WANG
Jianhua CHEN
Changsong ZHENG
School of Mechanical Engineering
Beijing Institute of Technology
5 South Zhongguancun Street, Haidian District
Beijing100081, China

E-mails: mabiao@bit.edu.cn,3120160206@bit.edu.cn, 3220190331@bit.edu.cn, 3220185033@bit.edu.cn, zhengchangsong@bit.edu.cn

Article citation info:
CERAN B, ORŁOWSKA A, KROCHMALNY K. The method of determining PEMFC fuel cell stack performance decrease rate based on the voltage-current characteristic shift. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 530–535, http://dx.doi.org/10.17531/ein.2020.3.16.

Bartosz CERAN
Agata ORŁOWSKA
Krystian KROCHMALNY

THE METHOD OF DETERMINING PEMFC FUEL CELL STACK PERFORMANCE DECREASE RATE BASED ON THE VOLTAGE-CURRENT CHARACTERISTIC SHIFT

METODA WYZNACZANIA SZYBKOŚCI SPADKU WYDAJNOŚCI STOSU OGNIW PALIWOWYCH TYPU PEMFC NA PODSTAWIE PRZESUNIĘCIA CHARAKTERYSTYKI NAPIĘCIOWO-PRĄDOWEJ*

The article presents mathematical model designed to estimate the rate of performance decrease in fuel cell stack. The fuel cell stack performance decrease rate is determined on the basis of stack average voltage measurements. The proposed model is used to determine power curve as well as exploitation indicators of fuel cell stack with a nominal power of 50 kW after 14 000 hours of continuous operation. The model is also used to determine the average voltage drop for the eleven-year fuel cell stack with a nominal power of 1,2 kW. In both studies, the values of exploitation indicators as well as their differences in relation to nominal values are determined.

Keywords: *fuel cells system exploitation, fuel cell stack exploitation indicators, stack performance decrease after years of operation, stack exploitation characteristics.*

Artykuł przedstawia model matematyczny przeznaczony do wyznaczenia szybkości spadku wydajności stosu ogniw paliwowych. Szybkość spadku wydajności stosu ogniw jest wyznaczana na podstawie wartości napięcia średniego stosu. Zaproponowany model wykorzystano do wyznaczenia krzywej mocy i wskaźników eksploatacyjnych stosu ogniw paliwowych o mocy nominalnej 50 kW po 14 000 h ciągłej pracy. Model wykorzystano także do wyznaczenia szybkości zmiany wartości napięcia średniego jedenastoletniego stosu ogniw paliwowych o mocy 1,2 kW. W obu badaniach wyznaczono wartości wskaźników eksploatacyjnych oraz ich różnice względem wielkości nominalnych.

Słowa kluczowe: *eksploatacja systemów ogniw paliwowych, wskaźniki eksploatacyjne stosu ogniw paliwowych, spadek wydajności stosu po latach eksploatacji, charakterystyka eksploatacyjna stosu.*

1. Introduction

Hydrogen fuel cells with Proton Exchange Membrane Fuel Cell (PEMFC) are considered one of the most promising and forward-looking technologies for generating electricity and heat. Their usage is foreseen for high-capacity power plants, small distributed sources [14] and transport sector [12, 30]. Fuel cells can be exploited in a wide range of electrical load variability, with favourable performance indicators such as: efficiency of the processing of fuel chemical energy into electricity, unit fuel chemical energy consumption, unit fuel consumption. Nevertheless, the reliability indicators of cell operation are still reaching unsatisfactory values and are the main reason for inhibiting the commercialisation of this technology on a large scale [21].

The improvement of reliability indicators and determination of operational indicators is now an up-to-date and important problem. The development of fuel cell technologies will be the main driver of the development of the hydrogen energy sector and hydrogen fuel-based transport. The research is being carried out on the use of hydrogen as an additional fuel for petrol and diesel engines [18, 20]. The analysis of the results of the study conducted by the authors of the work [20] showed the lack of appropriateness of using hydrogen as an additional fuel in compression ignition engines. Hydrogen power supply is mainly adapted by spark-ignition engines. However, the work [30] concluded that this is a temporary solution to pre-prepare and implement hydrogen storage and distribution infrastructure before the introduction of prospective fuel cells.

Due to continuously high costs of fuel cells, many studies of this technology are carried out using proprietary mathematical models [2, 34]. Work on modeling the fuel cell stack can be divided into two groups. The first includes works that present models designed to optimize stack parameters according to goal functions such as: minimizing construction costs, maximizing current density [1, 4, 6, 15, 27, 32].

In the article [4], the authors present a three-dimensional multiphase model of the PEM fuel cell designed to study the effect of assembly pressure on the contact resistance between the gas diffusion layer (GDL) and the bipolar plate. The correct selection of the mounting pressure allows you to make a stack with a longer life.

In the paper [6], the authors present a model designed to optimize fuel cell stack dimensions according to an improved version of the seagull optimization algorithm. The model allows to determine the operational characteristics of the stack, however, it does not take into account its displacement after the operational time. In turn, in the paper [1], the authors present a 500-watt fuel cell stack model with ion exchange polymer membrane implemented in Matlab/Simulink environment. The model is used to determine the reference value of electric current for any steady state.

In this paper [27], the authors present the processes taking place in the fuel cell stack and the developed numerical models aimed at mini-

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

mising the stack production costs and maximising current density. These models are designed to support the design of fuel cell stacks.

New techniques of fuel cell modeling are being developed, e.g. with the use of the so-called bonding diagrams [32]. Model tests are also carried out to increase energy efficiency by connecting a stack of cells with a thermoelectric generator into a hybrid system [15].

All the above mentioned mathematical models are built to support the design of the fuel cell stack. However, these models do not take into account the decrease in stack efficiency after years of operation.

The second group of publications on stack modelling include works that model the degradation processes taking place in the stack during its work. The authors focus on determining the influence of one of the structural elements, such as gas diffusion layer (GDL) [23], bipolar plates, polymer membrane [26] and catalyst layer on the electrodes of the stack [31], on the voltage-current characteristics. These models are used to minimize the source of degradation and increase the stack life.

The paper [26] presents a numerical model of fuel cells designed to test mechanical membrane degradation caused by local, non-pressurized membranes. In order to reduce the stresses, the authors suggest using an additional seal in connecting the electrode-electrolyte with GDL.

Subsequent works concern the determination of the probability of degradation of a given element [25, 35]. The authors of the article [25] proposed to use a foul three analysis tree to determine the probability of degradation of fuel cell stack elements. In turn, the article [35] proposes an innovative model of forecasting the degradation of a fuel cell with proton exchange membrane using the molecular filter and data extrapolation method.

Unfortunately, the above mentioned models and proposed methods are not oriented towards practical application because they do not show how the progressing degradation processes will affect the efficiency of the fuel cell stack after many hours or years of operation and, most of all, how the values of the performance indicators will change. The knowledge of changes in performance values will have a decisive impact on the correct determination of stack operating costs.

Publications on modelling hybrid power generation systems [16, 19, 28] confirm that the demand for a model showing the impact of fuel cell stack ageing on performance indicators is currently very high. The modelling techniques of hybrid systems presented in the above mentioned works do not take into account the decrease in stack efficiency after years of operation. Consequently, in such a hybrid system, energy imbalance may occur [24], which will result in increased consumption from the power grid [9].

Similarly, if a fuel cell is used in the transport sector [3, 11, 29], failure to take into account the decrease in stack efficiency may lead to an erroneous estimate of the vehicle's range or fuel demand, resulting in an erroneous estimate of the vehicle's operating costs.

In order to correctly determine the operating costs of a vehicle or a fuel cell hybrid electricity generation system, mathematical models should be built to determine the loss of performance of the cells after years of operation.

The main purpose of this article is to present a mathematical model which allows to determine the rate of decrease of fuel cell stack efficiency and the shift of the stack's operational characteristics in relation to the characteristics for catalog data after a given period of use. The model allows to determine the values of performance indicators, i.e. electrical efficiency, specific fuel chemical energy consumption for electricity production, specific fuel consumption for electricity production and to compare them with the values of indicators corresponding to the characteristics of the stack catalogues.

2. The mathematical model of the fuel cell

The mathematical models of fuel cells described in the literature make it possible to determine the operating characteristics for rated conditions (catalogue parameters). The proposed models do not take into account the decrease in stack efficiency after years of life and operation.

The analysis method proposed by the authors allows to determine the operational characteristics of the fuel cell stack for the rated data and the characteristics after "n" years of operation. The input variable of the model is, in this case, the rate of decrease of the average voltage value of one cell (one stack cell). On the other hand, the proposed algorithm can be used to determine the rate of drop in performance of the stack based on the shift between the catalogue characteristics and the stack characteristics after a given service life.

The method allows to quickly determine the efficiency and other performance indicators of the fuel cell stack.

The efficiency of converting the chemical energy of a fuel into electricity through a stack of fuel cells can be presented using formula (1) [7]:

$$\eta_{el} = \frac{P_{el}}{\dot{n}_{H_2} \cdot Q_{W_{H_2}}} = \frac{U \cdot I}{\dot{n}_{H_2} \cdot Q_{W_{H_2}}} \tag{1}$$

where: P_{el} – electrical power of the fuel cell stack, U – fuel cell stack voltage, I – electric current, $\dot{n}_{H_2}$ – hydrogen mole stream, $Q_{W_{H2}}$ – hydrogen calorific value related to 1 mole respectively.

The indicator of the unit chemical energy consumption of the fuel for the production of electricity through the stack is defined by formula (2):

$$q_{el} = \frac{\dot{n}_{H_2} \cdot Q_{W_{H_2}}}{P_{el}} \left[\frac{kJ}{kJ}\right] \tag{2}$$

The unit fuel consumption ratio for electricity production can be calculated using formula (3):

$$b_{el} = \frac{V_{H_2}}{P_{el}} \left[\frac{Nm^3 H_2}{kWh}\right] \tag{3}$$

where: V_{H_2} – hydrogen volume flow rate.

It should be noted that manufacturers of PEMFC fuel cell systems specify two efficiency values for converting the chemical energy of the fuel into electricity: the efficiency value related to the calorific value of hydrogen (Low Heating Value) and the efficiency value related to the heat of hydrogen combustion (High Heating Value).The analysis presented in this paper is carried out in relation to the calorific value of hydrogen, because the value of the q_{el} index is defined, in the classical theory of exploitation of generating sources, in relation to this value.

Defining the efficiency of a fuel cell in relation to its calorific value allows to compare its value with other technologies of electricity generation using the chemical energy of the fuel, e.g. conventional steam units, gas and gas-steam units, gas engines or technologies using biomass.

The hydrogen molar flux formula can be determined from the Faraday II electrolysis law (4):

$$q = n \cdot z \cdot F \tag{4}$$

where: q [C] – electric charge, n – number of gas moles, z [-] – number of electrons needed to release the molecule for H_2 = 2, i.e. 2 electron moles are needed to release 1 mole of H_2, for O_2 = 4, F [C/mol] – Faraday constant.

By dividing equation (4) by the time t, the relations to the electric current (5) is obtained:

$$I = \frac{n}{t} \cdot z \cdot F \tag{5}$$

From equation (5) the molar flow marked $\dot{n}$ (6) can be determined:

$$\dot{n} = \frac{n}{t} = \frac{I}{z \cdot F} \tag{6}$$

For a fuel cell stack, the hydrogen molar stream is proportional to the number of cells in the stack, hence:

$$\dot{n}_{H_2} = \frac{I \cdot n_{celek}}{z \cdot F} \tag{7}$$

where: n_{celek} – number of cells (single cells) from which the stack is built.

In the theory of fuel cells, a size called thermoneutral potential, which is defined according to formula (8) [22], applies in practice:

$$E_t^0 = -\frac{\Delta H^0_{H_2 0(g)}}{z \cdot F} \tag{8}$$

where: E_t^0 – thermoneutral potential [V], $\Delta H^0_{H_2 0(g)}$ – standard enthalpy of water formation in gaseous phase [kJ/mol], index “0” means standard conditions (T = 298 K, p = 10^5 Pa).

Thermalneutral potential is the theoretical value of the voltage that a fuel cell will reach with the theoretical assumption that 100 % of the supplied chemical energy stream will be converted into electricity.

The standard enthalpy of water formation in the gaseous phase ΔH¬H2O(g) energetically corresponds to the calorific value of hydrogen, assuming that water is the product in the gaseous state (9) [17]:

$$-\Delta H^0_{H_2 0(g)} = Q_{WH2} \tag{9}$$

By substituting equations (4), (8), (9) for equation (1), the formula for the efficiency of electricity generation by the fuel cell stack takes the form of formula (10):

$$\eta_{el} = \frac{U}{n_{celek} \cdot E_t^0} \tag{10}$$

The average voltage of a fuel cell stack is defined as the ratio of the voltage of the stack to the number of cells (single cells) in the stack:

$$U_{av} = \frac{U}{n_{celek}} \tag{11}$$

After substituting relation (11) to formula (10), the formula for stack efficiency can be presented with the help of the relation:

$$\eta_{el} = \frac{U_{av}}{E_t^0} \tag{12}$$

Formula (12) for the efficiency of converting the chemical energy of the fuel into electricity through the stack is very practical, because to determine the efficiency it is sufficient to measure the voltage of the stack and know the number of individual cells forming the stack. There is no need to measure the hydrogen flux used.

The change in stack efficiency is measured by changing the average stack voltage. The change in efficiency is, according to relation (13), directly proportional to the change in average stack voltage:

$$\Delta\eta_{el} = \frac{\Delta U_{av}}{E_t^0} \tag{13}$$

The change of average stack voltage is defined by formula (14):

$$\Delta U_{av} = U_{av} - \frac{dU_{av}}{dt} \cdot t \tag{14}$$

where: $\frac{dU_{av}}{dt}$ – the rate of change of average stack voltage over time [μV/h], t – fuel cell stack lifetime [h].

The power change of the stack can be determined on the basis of the relationship (15):

$$\Delta P_{el} = \Delta\eta_{el} \cdot \dot{n}_{H_2} \cdot Q_{WH_2} \tag{15}$$

The algorithm presented above allows to quickly determine the rate of change in the average voltage of the stack, and thus the rate of decrease in the generated electric power and the efficiency of converting the chemical energy of the fuel into electric power, as well as the increase in the values of the indicators of specific fuel consumption and specific chemical energy consumption of the stack.

The input data of the model are the catalogue parameters of the stack and the value of the stack voltage determined from the measurements. Based on formulas (1)-(15), the values of the stack’s operating indicators are determined and the operating characteristics are plotted for the catalogue parameters and after a given period of use. In order to determine the rate of decrease in the stack’s efficiency, the algorithm performs a series of simulations for different values of the rate of change of the average voltage of the stack with a given step and evaluates, with a given accuracy, the adjustment of the actual characteristics to the simulated one.

3. Determination of performance indicators based on the rate of change of average stack voltage values

The change of the average stack voltage ΔUav can be approximated by the linear function [33]. The rate of change in average voltage over time for systems of several tens of kilowatts is included, based on operational experience, in the range from about 3 to 9 μV/h [33]. This value depends on many factors, such as the culture of stack operation (observance of operating procedures, starting and stopping the stack), stack operating conditions (weather conditions, ambient temperature), the nature of stack operation (continuous, intermittent), etc.

A pilot plant built of 50 kW PEM cells (12 stacks of 4,2 kW) in the Netherlands, in Delfzijl, worked 14 000 hours without interruption.

Measurements during operation showed that the average stack voltage decreased at an average rate of 8 μV/h [33].

Based on equations 1 - 15 and the proprietary code developed in the Matlab environment, simulations were carried out to determine the power curve of the stack and the values of the performance indicators. The results of the simulation are shown in Figure 1 and Table 1.

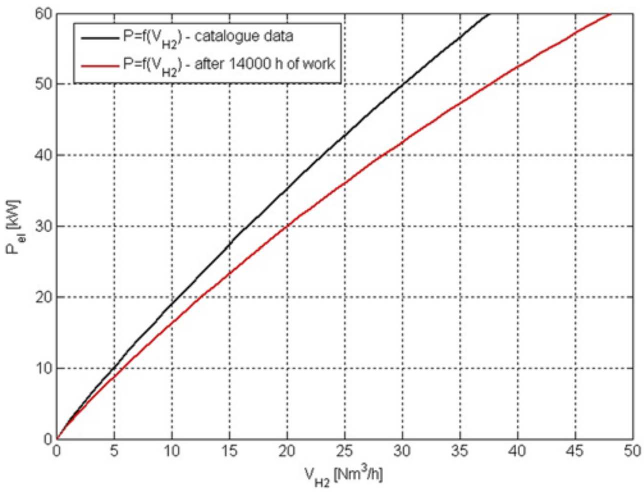


Fig. 1. Effect of U_{av} change on 50 kW fuel cell stack power curve - simulation tests

Table 1. Fuel cell stack performance indicators 50 kW - simulation tests

	Rated data	After 14 000 hours of operation
P [kW]	50	42.4
ΔP [kW]	−7.6	
η_{el} [-]	0.55	0.46
$\Delta\eta_{el}$ [-]	-0.09	
q_{el} [kJ/kJ]	1.81	2.16
Δq_{el} [kJ/kJ]	0.35	
b_{el} [Nm^3H_2/kWh]	0.60	0.72
Δb_{el} [Nm^3H_2/kWh]	0.12	

After 14,000 hours of continuous operation, the efficiency of the plant's electricity generation decreased by 16.36% from 55% to 46%. The effect of the decrease in efficiency is a decrease in the value of electrical power generated by the stack and an increase in the value of the indicators of specific chemical energy consumption of fuel and specific fuel consumption. The nominal capacity of the stack decreased by 15.2%. This means that the additional 7.6 kW of hydrogen chemical energy stream will be converted by the installation into thermal power at the expense of the value of generated electrical power. The value of the q_{el} index increased by 19.3% compared to the nominal value, while the value of the b_{el} index increased by 20%.

4. Determination of the rate of change of the average stack voltage from the measurement of the operating characteristics

In order to verify the model, it was used to match the voltage and current characteristics of the 1.2 kW fuel cell stack and to determine the rate of change of the average stack voltage. The tested fuel cell stack is an element of the NEXA training system, which is located in the energy conversion laboratory of the Wrocław University of Technology.

The NEXA system is a device designed for emergency power supply of both fixed and alternating current devices. Apart from the 1.2 kW stack, the system is made of [8]:

- hydrogen supply system – 20 MPa compressed hydrogen cylinders, hydrogen pressure regulator, hydrogen pressure regulator, pressure relief valve, solenoid valve to cut off fuel supply during system shutdown, hydrogen leakage detector,
- air supply system – a Roots blower,
- stack cooling system – the stack of cells in the NEXA system is cooled with air using a cooling fan,
- electronic control system – control computer, measuring sensors.

The age of the fuel cell stack is 11 years. Based on the simulation performed, the specific value of the average voltage change rate is 0,34 μV/h. For this value the best fit of the simulated and measured characteristics is obtained. The low value of the mean voltage change rate compared to a 50 kW system is due to the fact that the 1,2 kW cell stack consists of 46 cells. Additionally, the NEXA system operates in intermittent mode and is only used for research and teaching purposes.

Figure 2a shows the characteristics of the fuel cell stack: nominal, measurement-based characteristics (measured in March 2020) and simulation-based characteristics. Figure 2b shows the matching of measured and simulated characteristics.

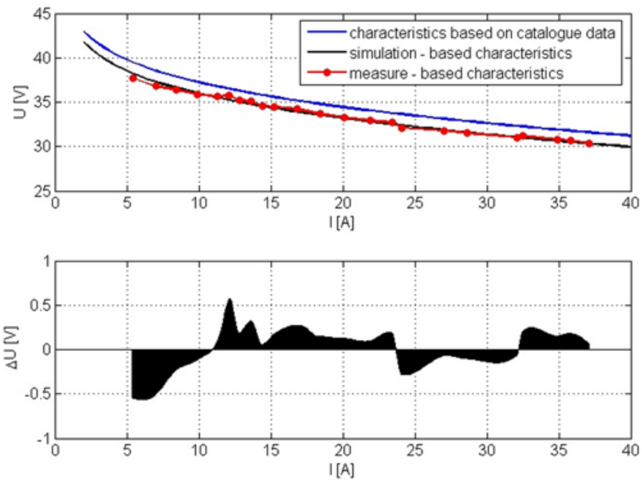


Fig. 2. a) Fuel cell stack characteristics: catalogue, simulated, measured, b) ΔU between simulated and measured characteristics

Deviations visible on the measurement characteristics are the effect of the anode flushing system during measurements [5]. At this point, the difference between the measured and simulated quantity is about 0.6 V. In most of the ohmic area [10, 13] of the voltage and current characteristics the difference between the points is about 0.2 V.

Table 2. Fuel cell stack performance indicators 1.2 kW

	Rated data	After 11 years
P_{el} [kW]	1.2	1.16
ΔP [kW]	−0.040	
η_{el} [-]	0.550	0.521
$\Delta\eta_{el}$ [-]	−0.029	
q_{el} [kJ/kJ]	1.840	1.920
Δq_{el} [kJ/kJ]	0.080	
b_{el} [Nm^3H_2/kWh]	0.610	0.640
Δb_{el} [Nm^3H_2/kWh]	0.030	

Table 2 shows a comparison of performance indicators calculated on the basis of catalogue data and simulation results. The low value of the average stack voltage change rate results in smaller changes in stack power and operating indicators. The nominal power of the fuel cell stack decreased from 1.2 kW to 1.16 kW after 11 years, reducing its value by 3.33%. The conversion efficiency of hydrogen chemical energy into electricity decreased by 5.27%. The fuel chemical energy conversion rate for electricity production increased by 4.35%. The specific fuel consumption for the production of one kWh of electricity increased by 4.92%.

Given the fact that the stack works under laboratory conditions with a high exploitation culture, the results should be considered as correct.

5. Summary

Updating the performance characteristics of fuel cells with ion exchange polymer membranes is a key issue in the development of hydrogen-based technologies, i.e. distributed generation of electricity and electromobility.

The proposed method allows to easily and quickly determine the decrease in efficiency of the fuel cell stack, which makes it possible to determine the current values of the performance indicators of the stack. On the other hand, the performance indicators determined after years of using the stack allow for a more accurate estimation of the operating costs of the fuel cell system.

The proposed method is a useful tool for carrying out a feasibility study of a project (e.g. hybrid power generation system with hydrogen storage, hydrogen car or hydrogen bus) on the basis of which the investor will be able to more accurately assess the risks associated with a given project by making technical and financial estimates a reality. The proposed model may facilitate the planning of long-term operation of fuel cell stacks in both distributed generation systems and hydrogen vehicles.

The advantages of the proposed method are its simplicity, short calculation time and the fact that it is sufficient to measure the voltage generated by the stack to determine the efficiency of the fuel cell stack.

The conducted simulations have shown that the value of the pile performance decrease depends on its nominal power (indirectly on the number of targets) and the mode and conditions of operation (continuous operation, intermittent operation). The model presented in the article contributes to the development of methods to update the performance characteristics of fuel cell stacks. These issues will be increasingly important in view of the expected decentralization of the electricity generation sector, the development of hybrid fuel cell stack generation systems and electromobility. The loss of fuel cell stack efficiency over the years of operation is an extremely important aspect for investors interested in new hydrogen technologies.

References

1. Abdelnasir O, Smith D, Alaswad A, Amiri A, Sodre JR, Lucchesi A. Proton-exchange membrane (PEM) fuel cell system mathematical modelling. Paper presented at SDEWES - 14th Conference on Sustainable Development of Energy, 2019; Water and Environment Systems, Dubrovnik, Croatia.
2. Abdul Rasheed RK, Liao Q, Caizhi Z, Chan SH. A review on modelling of high temperature proton exchange membrane fuel cells (HT-PEMFCs). International Journal of Hydrogen Energy 2017; 42(5): 3142-3165, https://doi.org/10.1016/j.ijhydene.2016.10.078.
3. Álvarez Fernández R, Corbera Caraballo S, Beltrán Cilleruelo F, Lozano JA. Fuel optimization strategy for hydrogen fuel cell range extender vehicles applying genetic algorithms. Renewable and Sustainable Energy Reviews 2018; 81(1): 655-668, https://doi.org/10.1016/j.rser.2017.08.047.
4. Atyabi SA, Afshari E, Wongwises S, Yan WM, Hadjadj A, Shadloo MS. Effects of assembly pressure on PEM fuel cell performance by taking into accounts electrical and thermal contact resistances. Energy 2019; 179: 490-501, https://doi.org/10.1016/j.energy.2019.05.031.
5. Barbir F. PEM Fuel Cell: Theory and Practice. New York: Elsevier Academic Press, 2005, 99-113.
6. Cao Y, Li Y, Zhang G, Jermsittiparsert K, Razmjooy N. Experimental modeling of PEM fuel cells using a new improved seagull optimization algorithm. Energy Reports 2019; 5: 1616-1625, https://doi.org/10.1016/j.egyr.2019.11.013.
7. Ceran B, Bernstein P.A. Operational characteristics of proton exchange membrane (PEM) fuel cells, Przegląd Elektrotechniczny 2014, 10: 102 - 105, 2014.
8. Ceran B, Długosz J, Kruczek-Pawlak H. Analiza energetyczna stosu ogniw paliwowych z jonowymienna membrana polimerowa PEMFC, Poznań University of Technology Academic Journals, 2016; 86: 301-312.
9. Ceran B, Orłowska A. The Impact of Power Source Performance Decrease in a PV/WT/FC Hybrid Power Generation System on the Result of a Multi-Criteria Analysis of Load Distribution. Energies 2019; 12(18): 3453, https://doi.org/10.3390/en12183453.
10. Ceran B. Charakterystyki eksploatacyjne stosu ogniw paliwowych typu PEMFC. Polityka Energetyczna - Energy Policy Journal 2014; 17(3): 135-146.
11. Chandrasekar C, Amruth Kumar L. A Novel Approach on Range Prediction of a Hydrogen Fuel Cell Electric Truck. SAE Technical Paper 2019; 2019-28-2514, https://doi.org/10.4271/2019-28-2514.
12. Chmielniak T, Lepszy S, Mońka P. Energetyka wodorowa - podstawowe problemy. Polityka Energetyczna - Energy Policy Journal 2017; 20(3): 55-66.
13. Dudek M, Celowski P, Lis B, Raźniak A, Dudek P. Laboratoryjny generator energii elektrycznej o mocy 360 W zawierający niskotemperaturowy stos ogniw paliwowych PEMFC chłodzony za pomocą medium ciekłego. Przegląd Elektrotechniczny 2016, 10: 235 - 242, 2014, https://doi.org/10.15199/48.2016.10.54.
14. Gharehpetian GB, Mohammad Mousavi Agah S. Distributed Generation Systems. Design, Operation and Grid Integration. Butterworth-Heinemann. An imprint of Elsevier, 2017, https://doi.org/10.1016/B978-0-12-804208-3.09993-3.
15. Guo X, Zhang H, Wang J, Zhao J, Wang F, Miao H, Yuan J, Hou S. A new hybrid system composed of high-temperature proton exchange fuel cell and two-stage thermoelectric generator with Thomson effect: Energy and exergy analyses. Energy 2020; 195: 117000, https://doi.org/10.1016/j.energy.2020.117000.
16. Hosseinalizadeh R, Shakouri GH, Amalnick MS, Taghipour P. Economic sizing of a hybrid (PV-WT-FC) renewable energy system (HRES) for stand-alone usages by an optimization-simulation model: Case study of Iran. Renewable and Sustainable Energy Reviews 2016; 54: 139-150, https://doi.org/10.1016/j.rser.2015.09.046.
17. Kabza A. Fuel Cell Formulary, www.kabza.de, dostęp: 31.03.2020.

18. Keršys A, Kalisinskas D, Pukalskas S, Vilkauskas A, Keršys R, Makaras R. Investigation of the influence of hydrogen used in internal combustion engines on exhaust emission. Eksploatacja I Niezawodnosc - Maintenance and Reliability 2013; 15 (4): 384-389.
19. Khan A, Javaid N. Optimum Sizing of PV-WT-FC-DG Hybrid Energy System using Teaching Learning-Based Optimization. International Conference on Frontiers of Information Technology (FIT) Islamabad, Pakistan, 2019; 1270-1275, https://doi.org/10.1109/FIT47737.2019.00033
20. Kruczyński S, Ślężak M, Gis W, Orliński P. Evaluation of the impact of combustion hydrogen addition on operating properties of self-ignition engine. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18 (3): 343-347, https://doi.org/10.17531/ein.2016.3.4.
21. Mayur M, Gerard M, Schott P, Bessler WG. Lifetime Prediction of a Polymer Electrolyte Membrane Fuel Cell under Automotive Load Cycling Using a Physically-Based Catalyst Degradation Model. Energies 2018; 11: 2054, https://doi.org/10.3390/en11082054.
22. O'Hayre R, Cha SW, Colella W, Prinz FB. Fuel Cell Fundamentals, 3rd Edition, Wiley 2016, https://doi.org/10.1002/9781119191766.
23. Park J, Oh H, Ha T, Lee YI, Min K. A review of the gas diffusion layer in proton exchange membrane fuel cells: Durability and degradation. Applied Energy 2015; 155: 866-880, https://doi.org/10.1016/j.apenergy.2015.06.068.
24. Paska J. Chosen aspects of electric power system reliability optimization. Eksploatacja I Niezawodnosc - Maintenance and Reliability 2013; 15 (2): 202-208.
25. Placca L, Kouta R. Fault tree analysis for PEM fuel cell degradation process modelling. International Journal of Hydrogen Energy 2011; 36(19): 12393-12405, https://doi.org/10.1016/j.ijhydene.2011.06.093.
26. Qiu D, Peng L, Liang P, Yi P, Lai X. Mechanical degradation of proton exchange membrane along the MEA frame in proton exchange membrane fuel cells. Energy 2018; 165: 210-222, https://doi.org/10.1016/j.energy.2018.09.136.
27. Secanell M, Jarauta A, Kosakian A, Sabharwal M, Zhou J. PEM Fuel Cells, Modeling. 2017; Springer, New York, NY, https://doi.org/10.1007/978-1-4939-2493-6_1019-1.
28. Shayeghi H, Shahryari E, Moradzadeh M, Siano P. A Survey on Microgrid Energy Management Considering Flexible Energy Sources. Energies 2019; 12(11): 2156, https://doi.org/10.3390/en12112156.
29. Sprik S, Thornton MJ, Brooks K, Tamburello DA. Performance Modeling of Materials-Based Hydrogen Storage Systems for Automotive Applications. 2017 AIChE Annual Meeting, 29 October - 3 November 2017, Minneapolis, Minnesota.
30. Sroka ZJ. Durability of engine components due to alternative fuels. Eksploatacja I Niezawodnosc - Maintenance and Reliability 2007; 4(36): 9-15.
31. Thangavelautham J. Degradation in PEM fuel cells and mitigation strategies using system design and control. T. Taner (Ed.), Proton Exchange Membrane Fuel Cell, London: Intech Open Ltd, 2018: 63-95, https://doi.org/10.5772/intechopen.72208
32. Vasilyev A, Andrews J, Jackson LM, Dunnett SJ, Davies B. Component-based modelling of PEM fuel cells with bond graphs. International Journal of Hydrogen Energy 2017; 42; 29406-29421, https://doi.org/10.1016/j.ijhydene.2017.09.004.
33. Verhage A, Gerits J, Manders T. Duration Tests of PEM Fuel Cells in a 50 kW Pilot Power Plant. In Proceedings of the 18th World Hydrogen Energy Conference (WHEC), Essen, Germany, 2010; 16-20 May: 63-67.
34. Wu H-W. A review of recent development: Transport and performance modeling of PEM fuel cells. Applied Energy 2016; 165: 81-106, https://doi.org/10.1016/j.apenergy.2015.12.075.
35. Zhou D, Wu Y, Gao F, Breaz E, Ravey A, Miraoui A. Degradation Prediction of PEM Fuel Cell Stack Based on Multiphysical Aging Model With Particle Filter Approach. IEEE Transactions on Industry Applications 2017; (53)4: 4041-4052, https://doi.org/10.1109/TIA.2017.2680406.

Bartosz CERAN
Agata ORŁOWSKA
Faculty of Environmental Engineering and Energy
Poznan University of Technology
ul. Piotrowo 3a, 60-965 Poznań, Poland

Krystian KROCHMALNY
Faculty of Mechanical and Power Engineering
Wrocław University of Science and Technology
ul. C.K. Norwida 1/3, 50-370 Wrocław, Poland

E-mails: bartosz.ceran@put.poznan.pl, agata.orlowska@put.poznan.pl,
krystian.krochmalny@pwr.edu.pl

Article citation info:
WALKER P, DOROSZUK B, KRÓL R. Analysis of ore flow through longitudinal belt conveyor transfer point. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 536–543, http://dx.doi.org/10.17531/ein.2020.3.17.

Piotr WALKER
Błażej DOROSZUK
Robert KRÓL

ANALYSIS OF ORE FLOW THROUGH LONGITUDINAL BELT CONVEYOR TRANSFER POINT

BADANIA SYMULACYJNE PRZESYPU WZDŁUŻNEGO PRZENOŚNIKA TAŚMOWEGO

A transfer point is an element of a belt conveyor prone to increased energy losses and to the risk of failure. It is also a location in which the receiving belt is particularly susceptible to damage. Except failure-free operation, a transfer point should offer minimal belt resistances to motion by ensuring that the transported material is placed centrally on the receiving belt, both spillage of the material and blockages are prevented, the process of particle defragmentation is limited, and also that noise and dust emissions to the environment are reduced. Ensuring that the above requirements are met requires inter alia the use of advanced simulation tests. The article analyzes the flow of ore particles stream through a longitudinal transfer point used in an underground copper ore mine. Discrete Element Method was used to identify the phenomena which occur while transferring ore onto the receiving conveyor. The research allowed key variables affecting the transfer point performance to be identified. It also resulted in a proposal of actions which can improve the performance of the transfer point and which are focused on saving energy and on minimizing the damage and wear of the receiving belt.

Keywords: *transfer point, Discrete Element Method, conveyor belt, abrasive wear, cumulative contact energy.*

Przesyp jest miejscem przenośnika taśmowego, w którym pojawia się ryzyko wystąpienia awarii, występują straty energii oraz może dochodzić do uszkodzenia taśmy odbierającej urobek. Poza bezawaryjnym funkcjonowaniem przesyp powinien dla zminimalizowania oporów ruchu taśmy zapewnić także centralne podawanie urobku, zapobiegać rozsypywaniu się transportowanego materiału, nie dopuszczać do powstawania zatorów, ograniczać proces defragmentacji urobku, a także minimalizować emisję hałasu oraz pyłów do otoczenia. Zapewnienie stawianych wymagań wiąże się z koniecznością stosowania m.in. zaawansowanych badań symulacyjnych. W artykule przeprowadzono analizę przepływu strugi urobku przez wybrany przesyp wzdłużny, stosowany w podziemnej kopalni rud miedzi. Przy użyciu metody elementów dyskretnych DEM dokonano oceny zjawisk zachodzących podczas przesypywania rudy na przenośnik odbierający, wskazano kluczowe zmienne opisujące jego pracę, a także zaproponowano działania udoskonalające pracę przesypu, zorientowane na zwiększenia jego energooszczędności oraz zmniejszenia negatywnego oddziaływania transportowanego nosiwa na taśmę przenośnika odbierającego.

Słowa kluczowe: *punkt przesypowy, Metoda Elementów Dyskretnych, taśma przenośnikowa, zużycie ścierne, skumulowana energia kontaktowa.*

1. Introduction

The function of the transfer point is to pass the transported material from the feeding conveyor to the receiving conveyor in such a manner as to prevent the material from being blocked [19]. In the case when a transfer point is badly designed, the transportation process may be interrupted, subjecting the mining company to considerable financial losses [12]. According to [10], downtime of the main belt conveyor in a copper mine in Chile, which produces 100,000 metric tons of ore per day, may lead to as much as $250,000 loss per hour. Except failure-free operation, a transfer point should offer minimal belt resistances to motion by ensuring that the transported material is placed centrally on the receiving belt, both spillage of the material and blockages are prevented, the process of particle defragmentation is limited, and also that noise and dust emissions to the environment are reduced [17].

Three transfer point design methods are currently used: analytical, simulational and experimental. The first analytical investigation of the trajectory of the material flowing through a transfer point was offered by C.E.M.A [1] in the 1960s. In 1988, Korzeń [20] presented an improved method for calculating the required hood curve (in hood and spoon transfer) and the particle flight path which allows calculations to be performed for both cohesive and non-cohesive materials. This topic was further researched by Roberts in 2003 [23]. In his calculations, Roberts allowed forthe friction between the particles and the elements of the transfer point, as well as for particle humidity. Over the following years, with increasing computational power available, transfer points started to be designed with the use of simulations based on Discrete Element Method (DEM), which consist in preparing a three-dimensional model of the investigated transfer point and in performing subsequent multi-variant analyses of how the design and operational parameters influence the transport of the material. In DEM simulations, forces acting on individual ore particles are determined at very short time intervals. In the first step, the algorithm identifies the total forces acting on the particles, and subsequently uses Newton's second law to find the acceleration and displacement for each of the

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

particles [18]. The two calculations are performed for one time step, and the algorithm proceeds to perform calculations for the next time step, using the data obtained in the previous step [14]. In order to simulate interactions between the particles, contact models are used.

Experimental, laboratory tests of transfer points are performed on specially adapted test stands, which employ high-speed cameras to identify the flow of material. The cameras allow an analysis of several tens to several hundreds of frames per second. Publication [11] describes an analysis of ore-flow through a transfer point, performed with the use of both analytical and simulation methods, and compares the effects with the experiment results. The research demonstrates that simulation methods have advantage over other methods. Publication [15], on the other hand, offers a comparative analysis of the analytical method proposed by Roberts [23] and of the DEM method, and indicates a high correlation between the results, while pointing to the limitations of the analytical methods.

The operation of the transfer points is related to the occurrence of concentrated resistances in the location where the material is fed [8]. Preliminary analyses performed for the KWB Bełchatów lignite mine suggest that these resistances may account for approximately 5% of the power loss on the conveyor [7]. The resistances occur when the material fed to the receiving conveyor has lower velocity tangential to the belt movement direction than the velocity of the belt itself, which indicates that the velocity of the material must be increased. Another component of the resistances is the friction of the material against the side skirts. Side skirts are used in the transfer point and installed over the belt of the receiving conveyor. They are located in the zone where the stream of material is formed and their function is to prevent individual grains from spilling outside the belt and to form the material into a shape appropriate for further transportation. As the material dropped on the receiving belt is typically characterized by turbulent movement, it should be stabilized before it leaves the skirts [3]. Additionally, during its flight, the transported material significantly increases its normal component of the velocity vector with respect to the conveyor plane. Therefore, when dropped on the belt, it may cause its damage and also negatively influence the condition of the feeding idler sets [21].

2. Simulation preparation

2.1. Calibration process

The effectiveness of the DEM simulations largely depends on the input parameters describing the simulated bulk material [2, 16]. The parameters used as the preliminary source of data for the calibration process of the investigated copper ore were obtained as part of own research [6]. The volumetric density was averaged on the basis of several samples, and the restitution coefficient was determined for steel, rubber and 3 types of rock present in the deposit. The ore-ore friction coefficient was determined with the use of a direct shear apparatus, and the ore-steel and ore-rubber coefficients were found with the use of an inclined plane. The rolling friction coefficient was assumed at 0.01, and other parameters were assumed on the basis of detailed literature studies, described in [6].

The test stand used in the experiment was similar in design to the stand described in [22]. It is a cubical box with a container holding the tested material in its upper part. The material slides on a plane which has variable inclination angle and which is located under the container. Upon releasing the lock, the sample stored in the container is allowed to leave it in the form of a stream of material which hits the plane and slides on it, eventually forming a fragment of a pile in the corner of the box. For the purpose of the tests, the device described in [22] was modified in such a manner that the material of the sliding plane could be modified. As a result, contact parameters were obtained for both steel and the carrying cover of the conveyor belt. In order to enable detailed observation of the behavior of the tested ma-

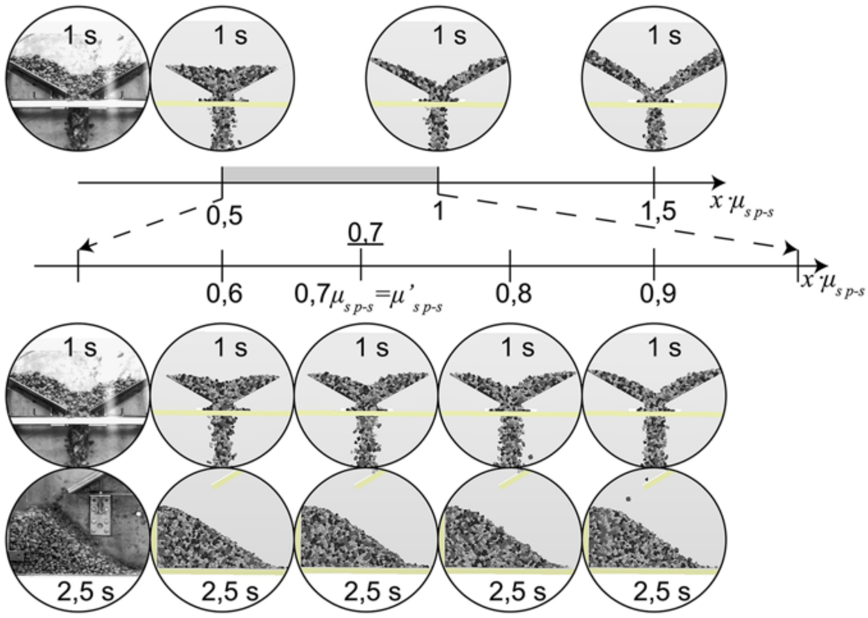


Fig. 1. Calibration step for particle-steel (p-s) coefficient of static friction (μ_s) – comparison of the characteristic parts of the simulation after 1 and 2.5 seconds for different multipliers (x) of friction coefficient

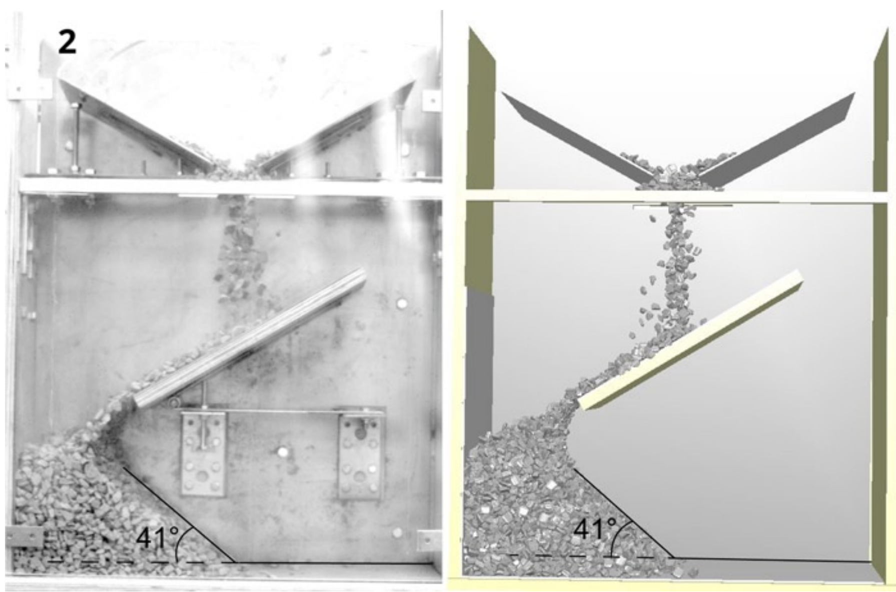


Fig. 2. On the left: a frame from the recorded experiment; on the right: the corresponding moment of the simulation

terial, the experiments were recorded with the use of Phantom Miro 120 high-speed camera, at 600 frames per second.

A number of simulations were performed, with each of the parameters modified in succession. The simulations were subsequently compared with the recorded experiments (Fig. 1). In the first step, the calibration process included the ore-ore contact parameters, and in the second stage – the ore-rubber parameters. As a result, a calibrated model was obtained for copper ore, which reflected the behavior of the actual material (Fig. 2) [5].

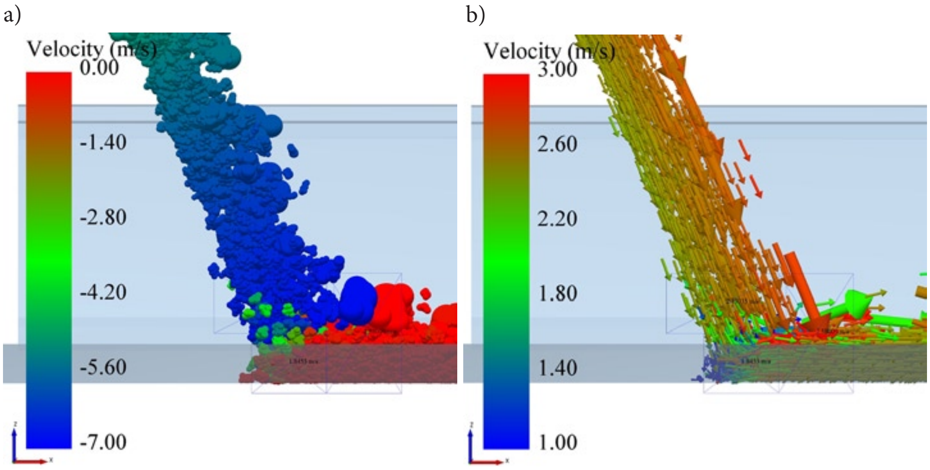


Fig. 4. View of the collision between the material and the belt during a DEM simulation: a) normal velocity of particles (recorded along the Z axis), b) tangential velocity of particles (along the X axis)

2.2. Preparation of the model

In the modeling process, the conveyor belt trough was filled with particles generated by the EDEM application. Their behavior was simulated with the use of the discrete element method and with allowance for the calibrated material parameters of the mined copper ore. Fig. 3 shows a simplified 3D model of the actual longitudinal transfer point operated in the mine. The transfer point comprised an adequately shaped upper (feeding) belt, lower (receiving) belt, feeding conveyor drum and side skirts. The geometry was modeled in the form of meshed surfaces with zero thickness. Only those surfaces were modeled which can be contacted directly by the transported material. This fact is due to the complexity of the calculation algorithm (and especially the detection of potential contacts between the material and the skirting elements), whose efficiency is also affected by the geometrical complexity of the model. The resulting geometry simulates the system to a satisfactory degree and is sufficient to perform DEM analyses.

In order to fully represent the grain size distribution, the SplitDesktop application was used to perform analyses on the basis of the collected photographic documentation of the material present on the top run of the feeding conveyor [24]. Only particles greater than 50 mm in diameter were modeled in the simulation, as the finer fraction would significantly increase the simulation preparation time, while its absence does not affect the results. Particles finer than 50 mm were cumulated and added to the finest generated fraction. In the simulation, the total assumed stream of the material mass at maximum capacity of the B1200 conveyor was 2072 Mg/h.

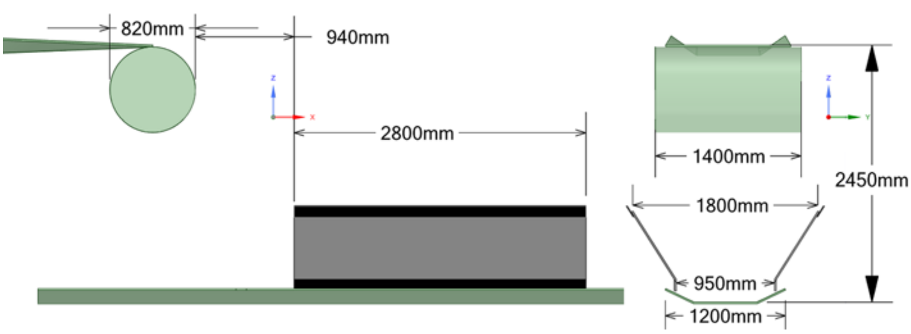


Fig. 3. Simplified 3D model of the analyzed transfer point used for the needs of the DEM simulation

3. Simulation

The criterion for the evaluation of the transfer point operation was based on component velocities of the material at the moment of collision with the belt. Known component velocities allow calculating abrasive wear of the belt at the location where the material accelerates. Material velocity tangential to the belt surface influences the energy-consumption of the transfer point and its concentrated resistances. When this velocity is lower than the velocity of the conveyor belt, losses occur [8, 9, 21]. However, velocity normal to the belt surface causes impact-related destruction of the carrying cover and of the feeding idler sets.

The analyzed design of the transfer point is a commonly used solution, in which the material is transferred directly from the belt of the feeding conveyor onto the belt of the receiving conveyor. Simulations demonstrated that at the moment of collision with the receiving belt, the stream of material has a normal velocity of approximately 6.3 m/s, which is further reduced so that in effect the entire energy of the falling material is received by the belt (Fig. 4a). In addition, virtual velocity sensors were added for individual particles in order to allow defining mean normal and tangential velocities of material particles during and after the impact.

In Fig. 4b, the vectors serve to present the analysis of particle behavior along axis X. Prior to the impact, the particles moved with a static velocity of approximately 2.37 m/s. At the moment of contact with the belt, the particles lost their velocity to 1.73 m/s and were subsequently accelerated, by the friction forces acting between the belt and the material, to the rated velocity of the belt (2.5 m/s). In order for the material to be accelerated, the friction force between the material and the belt must be greater than both the lift resistances of the material due to the inclination angle of the conveyor and the friction resistances of the material due to the side skirts [8]. This phenomenon accounts for the additional resistance to motion generated at the feed point.

The EDEM application was used to calculate the values of the energy received by the top run in the receiving conveyor over a 30-second period of its operation at rated capacity. For this purpose, the belt was divided into segments consisting of right triangles having sides 10 cm in length. Subsequently, the cumulated value of tangential and normal energy was calculated. The result served as a basis for identifying those belt fragments of the receiving conveyor which are most prone to wear. The simulations demonstrated that the greatest amount of energy is received by the middle fragment of the belt and by the side fragments. Less energy is received in the deflection zone on the idler set (Fig. 5).

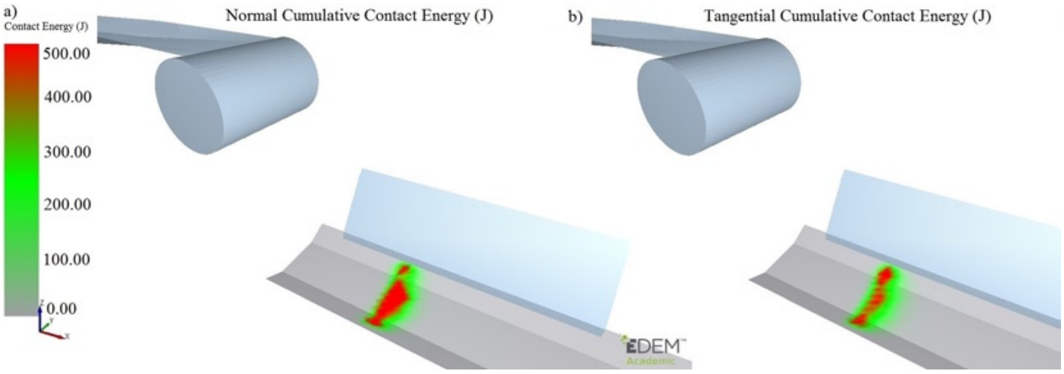


Fig. 5. Cumulative normal (a) and tangential (b) contact energy values determined in the EDEM software during the contact between the receiving conveyor belt and the material in the variant before modification

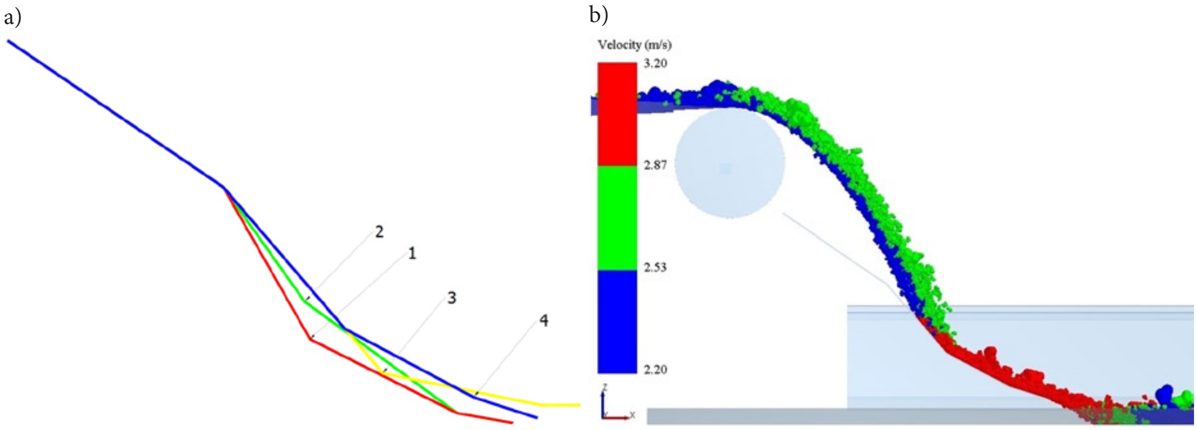


Fig. 6. Analyzed variants of the chute considered during the simulation tests (a) and the behavior of the material during the flows through chute No. 4 (b)

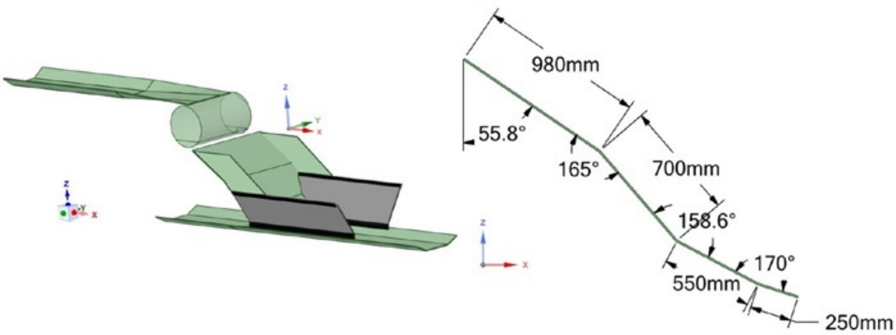


Fig. 7. Proposed design solution for longitudinal transfer chute

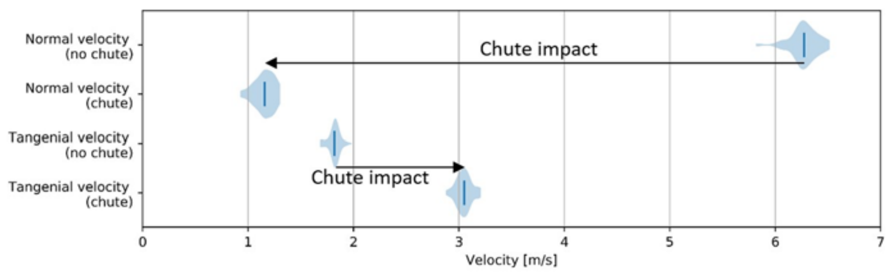


Fig. 8. Violin plot of tangential and normal velocity values recorded at the moment of the collision between the material and the belt for the tested variants of the longitudinal transfer point with indicated average velocity values

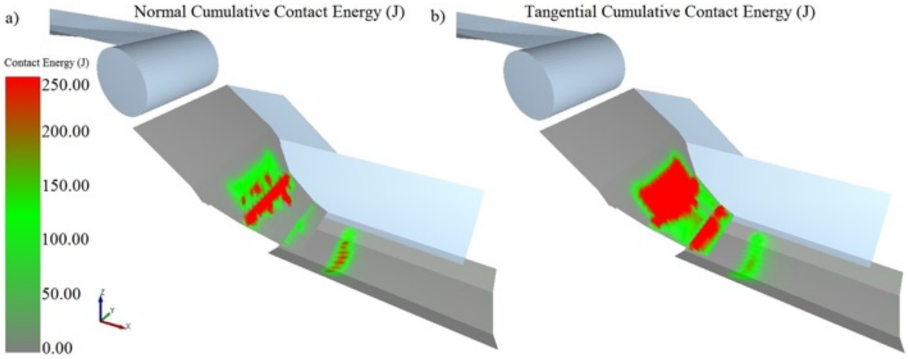


Fig. 9. Cumulated values of normal (a) and tangential energy (b) resulting from the contact of the transported material with the chute and with the receiving conveyor belt in the modified variant

As the analyzed transfer point is operated in a system of belt conveyors working at a constant speed of 2.5 m/s, the proposed optimizing solution was a chute. A chute appropriately adjusted to the discharge trajectory at the transfer point should reduce the value of the material component velocity normal to the belt and adjust its component tangential to the velocity of the receiving conveyor belt [9, 21].

The simulations were performed on four selected variants of the chute slope, with different numbers of plates and with their various inclination angles with respect to the vertical (Fig. 6a). The design differences in the tested solutions resulted from successive simulations of material behavior, during which the sensors available in the EDEM application were used to analyze the velocity components at the moment of material collision with the belt [13].

Fig. 6b shows changes of the tangential velocities of the material particles during the simulations for chute No. 4. Upon the collision with the receiving plate, the material was observed to increase its tangential velocity, which was then partly reduced due to friction against the lower section of the chute. As a result, the material acquired a velocity similar to the velocity of the receiving belt. Fig. 7 shows the geometric dimensions of the chute which provided the best optimization results.

Analogically to the analyses performed for the basic version, in the case of the accepted solution, both tangential and normal velocity components were also calculated for the material at the moment of collision with the belt. Fig. 8 shows a comparison of the obtained values. In the variant of the transfer point with a chute, the tangential velocity variable in time oscillates in the range from 2.8 m/s to 3.2 m/s, with mean velocity at approximately 3 m/s. This fact implies that the material is fed faster than the belt moves, and at the same time this value is sufficiently small not to cause the material to be spilled outside the transfer point. With the modernized version of the transfer point, the tangential velocity of material particles was increased by 1.2 m/s on average. Meanwhile, the calculated mean normal velocity of material particles in this variant is as much as 6-fold smaller and is approximately 1.15 m/s. This decrease allows a significant reduction of the destructive influence of the potential energy of the material masses falling on the conveyor belt and on the supporting idler sets.

Values of accumulated contact energy were used in order to identify the chute sections most susceptible to abrasive wear (Fig. 9). The sections of the chute most prone to damage due to normal and tangential forces have been indicated in red. Failure-free operation of the chute can be ensured by installing additional reinforcements in selected chute fragments (for example by using an appropriate construction material), and by frequently controlling their technical condition.

A comparison of the accumulated energies of the material acting on the belt allows an observation that the highest values of the accumulated normal energy occur in the case of the transfer point prior to modifications, and are 1028 J. The maximum value obtained for the variant with the chutez is 4-fold lower and is 265 J. The analysis of the accumulated contact energies from tangential forces also confirms a significant decrease for the variant of the transfer point with a chute (Fig. 5 and Fig. 9).

4. Analysis of abrasive wear of the belt on the receiving conveyor

Conveyor belt wear due to its abrasion caused during the feeding of the material may be calculated with the use of normal impact energy and the slip velocity between the material and the belt, as in the relationship below [9]:

$$W_a = \mu_b \cdot \rho_{bl} \cdot V_y^2 \left(V_b - V_x\right), \left[Pa \cdot \frac{m}{s}\right] \tag{1}$$

where:

- W_a – abrasive wear of the conveyor belt $\left[Pa \cdot \frac{m}{s}\right]$,
- μ_b – coefficient of static friction between the material and the conveyor belt [-],
- V_y – normal velocity of the material at the moment of impact against the belt [m/s],
- V_x – tangential velocity of the material at the moment of impact against the belt [m/s],
- V_b – belt speed [m/s],
- ρ_{bl} – volumetric density of the material [kg/m^3].

However, relationship (1) does not account for a number of factors which are important for defining belt wear at a certain point. Authors in publication [4] proposed to introduce such corrections to this equation that would allow for the inclination of the side fragments of the belt. Normal velocity of the material at the moment of impact against the belt V_y was substituted with velocity perpendicular to the belt $V_\perp$, which in the case of the inclined belt fragments is not equal to the velocity along axis Y. The second modification consisted in the introduction of the absolute value of the difference between belt velocity and material velocity parallel to the belt. This modification allowed the elimination of the negative values of belt wear in certain points selected for the analyses. Equation (1) thus transformed to be used in further calculations takes the following form:

$$W'_a = \mu_b \cdot \rho_{bl} \cdot V_\perp^2 \left|V_b - V_\parallel\right|, \left[Pa \cdot \frac{m}{s}\right] \tag{2}$$

where:

Table 1. Reference conditions for relative wear

Parameter	μ_{bs} [-]	ρ_{bls} [kg/m^3]	$V_{\perp s}$ [m/s]	V_{bs} [m/s]	$V_{\parallel s}$ [m/s]	W_s $[Pa \cdot m / s]$	ω_s [-]
Reference value	0.25	1000	1	2	1	250	1

$V_{\perp}$ – material velocity normal to the belt surface [m/s],

$V_{\parallel}$ – material velocity tangential to the belt [m/s].

The analyzed values obtained from the simulation data depend on the adopted sampling method. The simulations allow bin groups to be generated. These are virtual cubes situated over the analyzed conveyor belt. The bins record selected parameters of the material occurring in the simulation: its mean velocity or the masses of its particles. Importantly, belt wear is slower in the fragments filled with material to a lesser extent, i.e. the ones on which the material exerts a smaller force. Direct introduction of material weight into the equation would cause the calculated values of belt wear for the selected, differently sized bins to be incomparable. Therefore, standard values have been introduced (Table 1), which serve as reference for the values calculated in certain points. The proposed indicator of relative point wear (3) for each of the analyzed variants will allow the rate of abrasive wear to be determined for each point in comparison to the case in which the transported material has standard values referenced in Table 1, and the bin is completely filled with material:

$$\omega_{n,k} = \frac{\bar{F}_{\perp n,k}}{g \cdot V_{bin} \cdot \rho_{bls} \cdot \cos(\alpha)} \cdot \frac{W'_{a\,n,k}}{W_s}, [-] \tag{3}$$

where:

$\omega_{n,k}$ – indicator of point abrasive belt wear $[-]$,

$\bar{F}_{\perp n,k}$ – average force acting perpendicular to the belt on each measurement bin [N],

g – gravitational constant [m/s^2],

V_{bin} – volume of the bin in which the point measurement is performed [m^3],

ρ_{bls} – volumetric density of the reference material [kg/m^3],

α – belt inclination at a certain point [°].

If successive sampling points across the belt width are marked with the letter ‘k’, and along the belt length – with the letter ‘n’, and if the mean value of the point wear indicator is calculated for the identical n, then a relative wear indicator of the belt profile is formulated:

$$\vartheta_k = \frac{\sum_{n=1}^{n} \omega_{n,k}}{n} [-] \tag{4}$$

For the purpose of analyzing the abrasive belt wear, bins were added in the simulation. Their dimensions were 5x5 cm and 20 cm in height. A 10 s long simulation served to export data on mean velocity and the total mass of the particles inside the bins. The data were used to determine velocity distribution on the belt (Fig. 10). The particles which reach the belt in the variant of the transfer point with the chute clearly have a significantly higher initial velocity, and thus the zone in which they accelerate to the rated velocity is much shorter.

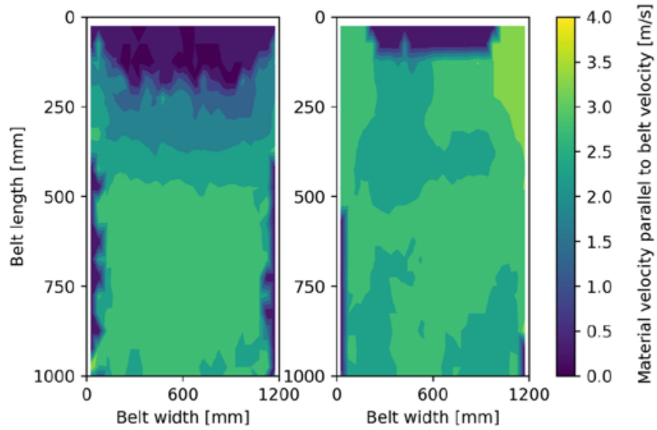


Fig. 10. Distribution of the material velocity tangential to the conveyor belt: on the left – without a chute, on the right – with a chute

The derived equations (3, 4) served to find belt areas most prone to abrasive wear (Fig. 11). The chute reduces belt wear, as it causes the lumps falling on the belt to have low normal velocity and high tangential velocity. Such a design solution allows point wear to be reduced by an order of magnitude. Provided in the graphs of Fig. 11, the belt profile wear indicators show mean values of the point wear indicator from the upper graphs, as observed along the belt width. The extremum observed in Fig. 11a was caused by the presence of larger lumps on the right side of the belt during the analyzed fragment of the simulation.

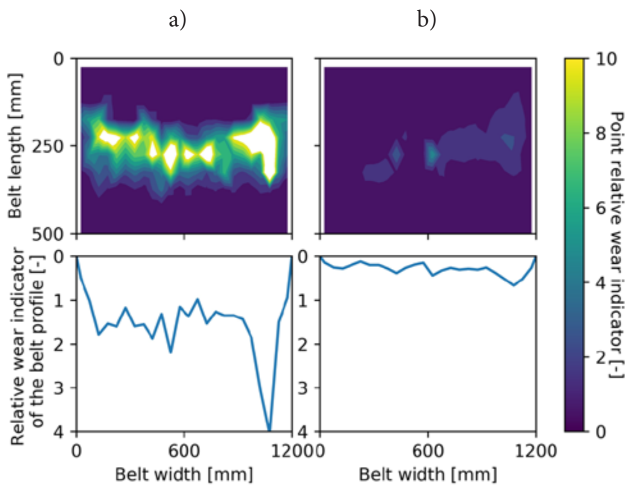


Fig. 11. Comparison of relative belt wear indicators: a) variant without chute, b) variant with chute

5. Conclusion

The article analyzes the flow of ore particles stream for a transfer point in a belt conveyor system operated in an underground copper ore mine. Discrete element method was used to identify the discharge trajectory and to evaluate the phenomena involved in the transfer of the material onto the receiving conveyor, with particular focus on the changing values of tangential and normal velocities of the material.

In order to fully simulate the behavior of the transported material in the model, a two-stage calibration test methodology was developed. It allowed the verification of the material parameters assumed in the simulations.

In the first stage of the simulation tests, the analysis covered the current solution, in which the feeding conveyor transfers the material directly onto the belt of the receiving conveyor. The stream of material was demonstrated to have a normal velocity of approximately 6.3 m/s at the moment of collision with the receiving belt. This velocity is further reduced so that in effect the entire energy of the falling material is received by the belt. The significantly reduced tangential velocity of the material particles at the moment of contact with the belt was observed to necessitate their acceleration until the belt rated velocity was reached. This acceleration generated additional resistances to motion.

In the second stage, optimization actions were suggested with a view to improving the energy consumption of the transfer point and to reducing the negative impact of the transported material on the belt of the receiving conveyor. The use of a modernized transfer point version with a chute resulted in a higher tangential velocity of the material particles and in a 6-fold reduction of its normal component. This solution significantly reduced the destructive influence of the material masses falling on the conveyor belt and on the feeding idler sets.

In order to identify the negative influence of the transported material on the belt of the receiving conveyor, a relative point wear indicator was proposed. Based on the results of the simulations, it allowed identifying the rate of abrasive wear individually for each of the analyzed belt fragments. The data from the simulations, read inside the bins, served to calculate the tangential velocity distribution for the material and to identify areas most prone to abrasive wear across the entire belt width. The analysis demonstrated that the use of a chute which redirects the material stream may result in as much as a 10-fold reduction of belt wear on the receiving conveyor.

Acknowledgement

The research work co-founded with the research subsidy of the Polish Ministry of Science and Higher Education granted for 2020.

Referenes

1. E.M.A, Belt Conveyors for Bulk Materials, Conveyor Equipment Manufacturers Association, 2005.
2. Coetzee C J. Calibration of the discrete element method. Powder Technology 2017; 310: 104-142, https://doi.org/10.1016/j.powtec.2017.01.015.
3. Czuba W, Furmanik K. Analiza ruchu ziarna w przesypie przenośników taśmowych. Transport Przemysłowy i Maszyny Robocze 2013; 4: 12-17.
4. Doroszuk B, Król R. Analysis of conveyor belt wear caused by material acceleration in transfer stations. Mining Science 2019; 26: 189-201, https://doi.org/10.37190/msc192615.
5. Doroszuk B. Badania Materiałowe i Modelowe Rudy Miedzi w Metodzie Elementów Dyskretnych (praca niepublikowana). Wrocław, 2019.
6. Doroszuk B, Walker P, Król R. Badania własności zróżnicowanej litologicznie rudy miedzi na potrzeby modelowania DEM. CUPRUM - Czasopismo Naukowo-Techniczne Górnictwa Rud 2019; 1: 5-19.
7. Gładysiewicz L et al. Optymalizacja rozwiązań technicznych przenośników taśmowych w PGE Bełchatów SA. Zadanie 3: Optymalizacja przesypów przenośnikowych. Raporty Inst. Gór. PWroc. Ser. SPR nr 1, 97 s. (praca niepublikowana), Wrocław, 2011.
8. Gładysiewicz L, Przenośniki Taśmowe teoria i obliczenia, Wrocław: Oficyna wydawnicza Politechniki Wrocławskiej, 2003.
9. Grima A, Hastie D, Fraser T, Wypych P. Discrete Element Modelling: Trouble-Shooting And Optimisation Tool For Chute Design. BELTCON 16 International Materials Handling Conference. Johannesburg, 2011.
10. Gutierrez A, Garate G A. Design of a Chute for Multiple Operating Conditions. ASME International Mechanical Engineering Congress and Exposition, 2014, https://doi.org/10.1115/IMECE2014-36414.
11. Hastie D, Wypych P. Experimental validation of particle flow through conveyor transfer hoods via continuum and discrete element methods. Mechanics of Materials 2010; 4: 383-394, https://doi.org/10.1016/j.mechmat.2009.11.007.
12. Hastie D. Belt Conveyor Transfers: Quantifying and Modelling Mechanisms of Particle Flow. Wollongong: University of Wollongong, 2010.
13. Hastie D, Grima A, Wypych P. Modelling and Design of Complete Conveyor Transfers. 2nd Annual Conveyors in Mining, Perth , 2008.
14. Huque S T. Analytical and Numerical Investigation Into Belt Conveyors Transfers, Wollongong: University of Wollongong, 2004.
15. Ilic D, Mcbride B, Katterfeld A Validation of continuum methods utilizing discrete element simulations as applied to a slewing stacker transfer chute. 9th International Conference on Bulk Materials Storage, Handling and Transportation. Newcastle, 2007.
16. Karwat B, Machnik R, Niedźwiedzki J, Nogaj M, Rubacha P, Stańczyk E. Calibration of bulk material model in Discrete Element Method on example of perlite D18-DN. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (2): 351-357, https://doi.org/10.17531/ein.2019.2.20.
17. Katterfeld A, Groger T, Hachmann M, Becker G. Usage of DEM simulations for the development of a new chute design in underground mining. 6th International Conference for Conveying and Handling of Particulate Solids. Brisbane, 2009.
18. Katterfeld A. Understanding granular media: from fundamentals and simulations to industrial application. Granular Matter 2017; 21, https://doi.org/10.1007/s10035-017-0765-y.
19. Kessler F, Prenner M. DEM - Simulation of Conveyor Transfer. FME Transactions 2009: 185-192.
20. Korzeń Z. The Dynamics of Bulk Solids Flow on Impact Plates of Belt Conveyor Systems. Bulk Solids Handling 1988; 12: 689-697.
21. Mascarenhas F, Mesquita A, Mesquita A L. Simulation of transfer chute operation using the Discrete Element Method. Cilmace 2013. Pirenópolis, Brazil, 2013.
22. Quist J, Evertsson M. Framework for dem model calibration and validation. Proc. of the 14th European Symposium on Comminution and Classification, Gothenburg, Sweden, 2015.

23. Roberts A. Chute Performance and Design for Rapid Flow Conditions. Chemical Engineering and Technology 2003: 163-170, https://doi.org/10.1002/ceat.200390024.
24. Walker P, Kawalec W, Król R. Application of the Discrete Element Method (DEM) for Simulation of the Ore Flow Inside the Shaft Ore Bunker in the Underground Copper Ore Mine. Advances in Intelligent Systems and Computing 2019; 835: 633-644, https://doi.org/10.1007/978-3-319-97490-3_60.

Piotr WALKER
Błażej DOROSZUK
Faculty of Geoengineering, Mining and Geology
Wroclaw University of Science and Technology
Wybrzeże Stanisława Wyspiańskiego 27, 50-370 Wrocław, Poland

Robert KRÓL
Department of Mining and Geodesy
Faculty of Geoengineering, Mining and Geology
Wroclaw University of Science and Technology
Wybrzeże Stanisława Wyspiańskiego 27, 50-370 Wrocław, Poland

E-mail: piotr.a.walker@gmail.com, blazej.doroszuk@pwr.edu.pl, robert.krol@pwr.edu.pl

Article citation info:
KARABACAK YE, GÜRSEL ÖZMEN N, GÜMÜŞEL L. Worm gear condition monitoring and fault detection from thermal images via deep learning method. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 544–556, http://dx.doi.org/10.17531/ein.2020.3.18.

Yunus Emre KARABACAK
Nurhan GÜRSEL ÖZMEN
Levent GÜMÜŞEL

WORM GEAR CONDITION MONITORING AND FAULT DETECTION FROM THERMAL IMAGES VIA DEEP LEARNING METHOD

MONITOROWANIE STANU I WYKRYWANIE BŁĘDÓW PRZEKŁADNI ŚLIMAKOWEJ NA PODSTAWIE TERMOGRAMÓW Z WYKORZYSTANIEM METODY GŁĘBOKIEGO UCZENIA

Worm gearboxes (WG) are often preferred, because of their high torque, quickly reducing speed capacity and good meshing effectiveness, in many industrial applications. However, WGs may face with some serious problems like high temperature at the speed reducer, gear wearing, pitting, scoring, fractures and damages. In order to prevent any damage, loss of time and money, it is an important issue to detect and classify the faults of WGs and develop the maintenance plans accordingly. The present study addresses the application of the deep learning method, convolutional neural network (CNN), in the field of thermal imaging that were gathered from a test rig operating on different loads and speeds. Deep learning approaches, have proven their powerful capability to exploit faulty information from big data and make intelligently diagnostic decisions. Studies concerning the condition monitoring of WGs in the literature are limited. This is the first study on WGs with infrared thermography rather than vibration and sound measurements which have some deficiencies about hardware requirements, restricted measurement abilities and noisy signals. For comparison, CNN was also trained, with vibration and sound data which were collected from the healthy and faulty WGs. The results of fault diagnosis show that thermal image based CNN model on WG has achieved 100% success rate whereas the vibration performance was 83.3 % and sound performance was 81.7%. As a result, thermal image based CNN model showed a better diagnosing performance than the others for WGs. Moreover, condition monitoring of WGs, can be performed correctly with less measurement costs via thermal imaging methods.

***Keywords**: fault diagnosis, worm gears, thermal imaging, convolutional neural networks, GoogLeNet, condition monitoring.*

W wielu zastosowaniach przemysłowych preferuje się przekładnie ślimakowe, ze względu na ich wysoki moment obrotowy, możliwość szybkiej redukcji prędkości i dobrą sprawność zazębienia. Jednakże przekładnie tego typu narażone są często na poważne problemy, takie jak wysoka temperatura przy reduktorze prędkości czy też zużycie, pitting (wżery), zatarcie, pęknięcie lub uszkodzenie kół zębatych. Zapobiec takim uszkodzeniom, i związanym z nimi stratom finansowym i czasowym, można poprzez wykrywanie i klasyfikowanie błędów przekładni i odpowiednie opracowanie planów konserwacji. Niniejsze badanie dotyczy zastosowania metody głębokiego uczenia oraz splotowych sieci neuronowych (SSN) do monitoringu stanu przekładni na podstawie termogramów zarejestrowanych na stanowisku testowym pracującym przy różnych obciążeniach i prędkościach. Podejścia oparte na uczeniu głębokim umożliwiają efektywne wykorzystanie informacji o błędach pochodzących z dużych zbiorów danych i podejmowanie trafnych decyzji diagnostycznych. Niewiele z dostępnych publikacji poświęconych jest monitorowaniu stanu przekładni ślimakowych. Niniejsza praca jako pierwsza przedstawia badania przekładni ślimakowej z zastosowaniem termografii zamiast zwyczajowo prowadzonych pomiarów drgań i dźwięku, które mają pewne wady dotyczące wymagań sprzętowych, ograniczonych możliwości pomiarowych i głośności sygnałów. SNN opartą na danych termicznych porównano z siecią, którą uczono na zbiorach danych wibracyjnych i akustycznych pochodzących z prawidłowo działających i uszkodzonych przekładni ślimakowych. Wyniki diagnostyki uszkodzeń pokazują, że model SSN przekładni ślimakowej oparty na obrazie termicznym osiągnął stuprocentową (100%) skuteczność, podczas gdy skuteczność modeli opartych na danych wibracyjnych i akustycznych wyniosła, odpowiednio, 83,3% i 81,7%. Tym samym, model SNN oparty na obrazie termicznym pozwalał na trafniejsze diagnozowanie przekładni ślimakowej niż pozostałe modele. Ponadto zastosowanie metod opartych na termografii pozwala na poprawne monitorowanie stanu przy niższych kosztach pomiaru.

***Słowa kluczowe**: Diagnostyka błędów, przekładnie ślimakowe, termografia, splotowe sieci neuronowe, GoogLeNet, monitorowanie stanu.*

1. Introduction

Industrial condition monitoring applications have improved a lot depending on the novel monitoring technologies and artificial intelligence decision making methods. In places where production continues uninterrupted such as factories, power generation facilities, lifting appliances, elevators, it is extremely important for systems to operate without any faults in terms of cost and work safety [31]. Due to the increased complexity and precision of these systems and engineering applications, condition monitoring and reliability problems become more prominent [41]. During condition monitoring, the system is real time observed and real time measurements are done. By interpreting the measurements, whether there is a fault in the machine or not, is determined and the maintenance plan is applied accordingly. Establishing a reliable condition monitoring system, especially for gear faults, ensures a smooth operation during the working of many industrial equipment [33].

Worm gear condition monitoring studies also plays a critical role in the maintenance plan since they are commonly used for power and motion transmission in many industrial applications and machines. They contain a mechanism consisting of worm screw and worm wheel that work together and differ from other types of reducers due to their lightness, simplicity and high gear ratio [11]. The main problem with a WG is how it transfers power. The spiral motion allows huge amounts of reduction in a comparatively small amount of space however this motion can also cause a problematic condition that is called sliding wear. During the working of a WG set, as the worm slides across the apex of the wheel, it slowly rubs off the lubricant film, and as a result, the worm rubs at the metal of the wheel in a boundary lubrication regime. When the worm surface leaves the wheel surface, it picks up more lubricant, and starts the process over again on the next revolution. This contact between the worm and the wheel in less lubricant conditions can cause wear and high temperature which need to be diagnosed. The most common problems faced with WGs are the high temperature at the speed reducer causing oil leakage, gear wearing, pitting, scoring and bearing fractures and damages. In order to prevent any damage, loss of time and money, serious precautions should be taken or predictive maintenance methods should be applied.

There are commonly used techniques such as vibration monitoring, acoustic monitoring, wear debris analysis, motor current analysis, in the literature for monitoring the conditions of gearboxes. Ghodake et al. [8] reviewed fault detection studies for worm gearboxes. Goyal et al. [9] performed a review study for condition monitoring and fault diagnosis for fixed axis gearboxes. Carden and Fanning [5] reviewed condition monitoring studies based on vibration analysis. Liu and Zhang [26] put ahead failure modes, condition monitoring and fault diagnosis methods for wind turbine bearings. Sait and Sharaf-Eldeen [35] conducted a review study on fault diagnostic and prognostic with the vibration analysis technique. Sharma et al. [37] reviewed different statistical condition indicators in time and frequency domains for gear faults. Lei et al. [18] reviewed studies of empirical mode decomposition method to diagnose faults of rotating machinery. Osman and Wang [32] proposed an improved Hilbert–Huang transform technique to diagnose faults of bearings using vibration signal analysis. They verified the effectiveness of the suggested technique in feature extraction and analysis by experimental tests. Wang et al. [45] proposed a hybrid technique by using complex wavelet transform for health diagnosis of rotating machines. They used numerical simulation and experimental studies under varying operating conditions to show the effectiveness of the hybrid technique. Loutas et al. [28] used on-line signal measurements and studied condition monitoring of a single-stage gearbox having artificial cracks. They used acoustic emission and vibration measurements to make use of this purpose. Zhang et al. [48] improved a new feature extraction method named Narrow-band Interference Cancellation to diagnose gear faults easier. Zhang and Zhao [49] proposed a compound fault detection approach based on time synchronous resampling and adaptive variational mode decomposition for gearboxes. With the experimental data analysis, they showed that the method is valid and can be used for fault detection of gearbox. Zhang et al. [50] proposed a fault diagnosis method based on singular value decomposition and radial basis function neural network to detect weak gear fault signals. Zhao et al. [51] proposed a local feature-based recurrent unit networks for monitoring health of machines. The method is a hybrid approach that integrates feature design with automatic feature learning.

Vibration monitoring has been considered as the most prevalent technique because it is easier to gather vibration data with vibration sensors [36]. Vibration sensors are versatile tools that measure acceleration for various applications [43]. However, the signals obtained from vibration sensors are very prone to the position of placements, harsh working conditions and high temperature [10]. Vibration and acoustic measurements were also conducted together to improve the performance [27]. Thermal Imaging Technique (TIT) is introduced as a contactless, continuous and easy to implement technique [3] that can sense the radiation in a long-infrared range (9-14μm) of the electromagnetic spectrum and produce thermal images known as thermograms [29]. A single thermogram recorded in a very short time interval can contain multiple temperature points and provides information about the system being studied [34]. In addition, infrared thermography technique is used as a condition monitoring tool in which contactless and real-time abnormal temperature distribution can be detected [3]. Singh et al. [39] presented a comparative fault detection approach on asynchronous motors using the infrared thermography. They also detected in-turn and cooling system failures of induction motor with thermal imaging techniques [40]. Wakar and Demetgül [46] collected thermal images for normal and faulty conditions under different speeds through an experimental system containing worm gearbox, and a multilayer perceptron was developed based on vibration and acoustic emission signals. Janssens et al. [13] detected eight different faults of bearings using thermal image based fault detection system and proposed two new features for thermal imaging of rotating machines. Al-Musawi et al. [2] developed a new coloring model and classified different bearing faults of a three-phase induction motor based on thermal image segmentation. Younus and Yang [47] developed a new intelligent detection system to classify different operating conditions based on the support vector machine and linear discriminant analysis using data from infrared thermography. Zhang et al. [50] used singular value decomposition and Radial Basis Function neural network for the detection of gear faults. In different condition monitoring studies, artificial neural network (ANN) based methods and computational modelling methods were used [22].

Despite the use of many fault detection and condition monitoring studies with the use of numerous diagnosing algorithms, such as support vector machines, empirical mode decomposition, Wigner-Ville distribution, short-time Fourier Transform, etc [22], as the evaluation data become larger, an efficient deep learning method, CNN, was introduced, in which extracting special features from the entire data is not needed. Nowadays, condition monitoring of mechanical faults in rotating components, are being prescribed by intelligent diagnosing methods such as artificial intelligence techniques and deep learning [9]. Deep learning is one of the newest machine learning techniques and has recently been used in condition monitoring studies and has been proved to be an effective methodology to improve the safety and reliability of gearboxes [17]. Zhao et al. [51] was the first who redefined the representation learning of raw data with the deep learning method. They proposed a local feature-based gated recurrent unit network for fault diagnosis. Li et al. [25] presented a modified deep learning method that can be used in case of limited data access and compared the method with traditional deep learning methods in terms of diagnostic success using datasets of two different gearboxes. Li et al. [22] proposed a new method named augmented convolution sparse auto-encoder to diagnose pitting faults of gears by using acoustic emission signals. Comparative analysis shows that the proposed method gives better results than fully-connected layer neural network and convolutional neural network. Few researches were carried out fault detection and classification with thermal image based deep learning methods. Li et al. [24] have used thermal imaging as a cost-effective and practical predictive maintenance method. In their study, they processed thermal images with a convolutional neural network (CNN) and allowed online remote condition monitoring of a gearbox. Their proposed method predicted faults on the gearbox with higher accuracy rate than their vibration-based counterparts. Janssens et al. [14] analyzed infrared thermal video images with deep learning algorithm to determine the condition of the machine automatically. Then, with deep learning, they detected the machine's faults and oil level. Li et al. [23] proposed a CNN method based on infrared thermal images to detect faults of rotors and bearings on rotating machines and compared the results with conventional vibration or sound based approaches and other CNN types. Condition monitoring studies based on thermal images with CNN are limited to these publications. However, many studies have been carried out for fault detection with vibration and sound based deep learning. Jing et al. [15] proposed a CNN based feature learning and fault diagnosis method for gearboxes and showed that the CNN is more appropriate to learn features from vibration signal. Li et al. [19] proposed a deep random forest fusion method by using acoustic and vibration signals. They showed that the deep learning fusion can develop the ability of fault detection and diagnosis for gearbox. In another study, Li et al. [20] studied deep statistical feature learning to diagnose rotating machinery faults. The experiments show that the proposed method based on deep learning has the best fault classification rate and potential for diagnosing of rotating machinery faults.

This study presents a condition monitoring application with CNN, based on thermal images of a WG test rig which simulates different operating loads and speeds. The novelty of this study lies in the application of CNN for detecting faults on WGs from thermal imaging data. Vibration and sound signal measurement and analysis devices are relatively expensive than thermal cameras. Moreover, compared to signal analysing devices, a camera can examine the entire surface of WG. The operating speed and loading rate have a major impact on the gearbox temperature, sound and vibration behaviour. In order to distinguish the fault and the effect of working conditions, measurements should be done for all of the operating conditions. For this reason, in this study a test rig that can work at different speeds and loading rates was used to collect thermal images, vibration and sound signals from healthy and faulty WGs. Since an industrial WG is loaded at different rates and operated at different speeds in the real working environment, measurements at constant speed and load are only valid for a limited operating range. For comparison reasons, vibration and acoustic data were also gathered for all working conditions. The effects of varying resolution on performance of CNN are also depicted. As far as the authors' search, the results of the present study are of vital importance to the development of WG knowledge and industry.

The rest of this paper is formed as follows. Section 2 describes the experimental system, test rig and measurement elements. Section 3 describes the methodology and theoretical background of CNN-based fault diagnosis and proposed models for WG. Section 4 describes experimental validation and performance comparison with proposed models. Finally, section 5 gives conclusions.

2. Experimental System

2.1. Test Rig

The test rig shown in Fig. 1 was built to verify the CNN-based fault diagnosis method proposed for WG. The test rig consists of three main units: control panel, loading device and single stage WG connected to an induction motor. The loading device, control panel and WG are mounted on a steel platform. The induction motor is driven by an inverter and coupled to the WG. An elastic coupling is used between the output shaft of the WG and an electromagnetic powder brake is the loading device. A radial fan protects the loading device from excessive heat.

The inverter is adjusted by a digital panel placed on the control panel. The DC loading device is powered by a transformer which has 5 different outputs so that the system can be loaded at different rates as 0%, 25%, 50%, 75% and 100%. Real working conditions are simulated with the inverter and the loading device powered by a multiple output transformer. The output speed of the gearbox is measured with a tachometer and the input voltage of the loading device with an AC/DC voltmeter.

2.2. Measurement Elements

Three different measuring devices were used in the experiments: thermal camera, accelerometer and microphone. The TESTO 880 thermal camera with 9 Hz image refreshing frequency (Fig. 2 (a)). Thermal sensitivity of the camera is less than 0.1 °C and it can focus manually up to a distance of 10 cm. It has 160x120 pixel detector and 32°x24° standard lens. The second one is a PCB ICP® type piezoelectric accelerometer for vibration data. (Fig. 2 (b)). The accelerometer has a sensitivity of 1.02 mV/(m/s²) and allows measurements in the frequency range of 0.5-10000 Hz. Phantom powered Behringer ECM8000 condenser microphone is used for sound measurements (Fig. 2 (c)). The microphone has 600-ohm resistance and −60 dB sensitivity within the frequency range of 15 Hz - 20 kHz. In addition, m + p VibPilot dynamic signal analyser with 4 analog outputs was used to collect and monitor data with accelerometer and microphone.

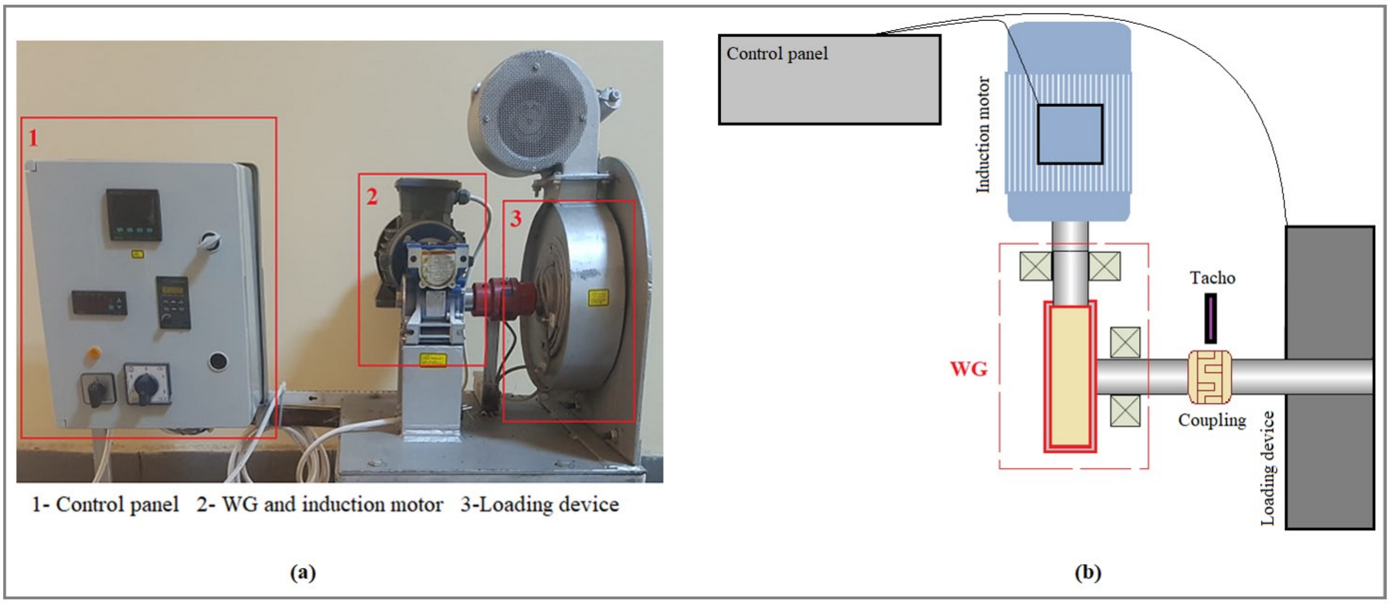


Fig. 1. Experimental setup: (a) test rig (b) the schematic diagram

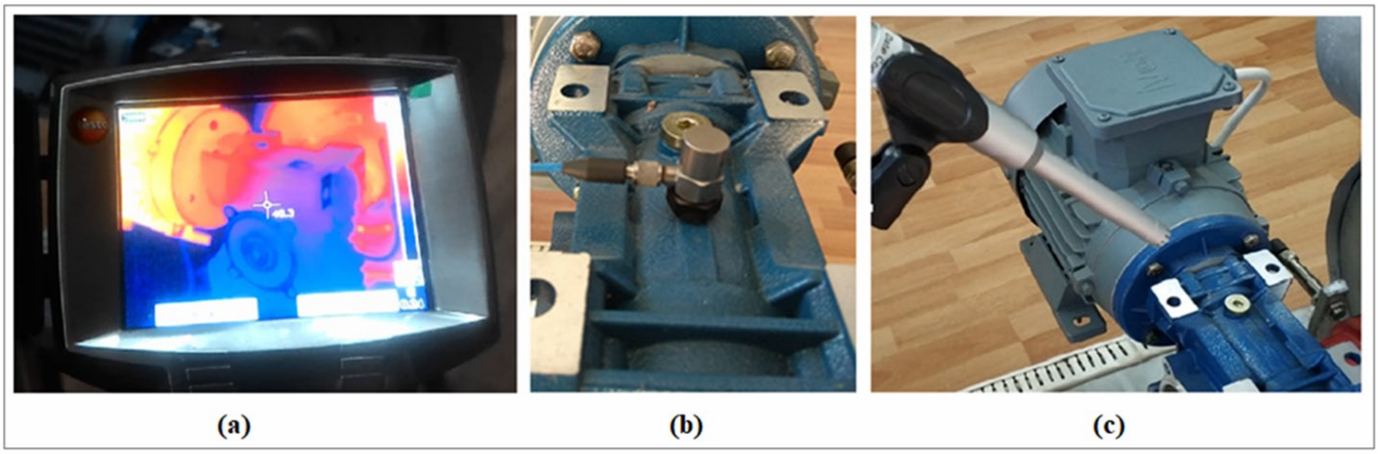


Fig. 2. (a) thermal camera, (b) accelerometer, (c) microphone

3. Methodology

This section presents the methodology used to classify WG faults that are healthy, wear, pitting and tooth breakage. Deep learning, one of the recent machine learning methods, is used as the main tool. As stated in previous works [22], deep learning do not require manual extraction of fault features and also achieve better fault detection results. It has the ability to reduce the number of network parameters through local weight sharing and to avoid the over-fitting of the network when the number of samples is insufficient.

3.1. Theoretical Background of CNN

Convolutional neural network (CNN) is one of the most used types of deep learning. CNN consists of convolutional and subsampling layers. Each of these layers has a certain topographic structure and the layers contain a different set of neurons. Each neuron is also linked to neurons in the previous layers. Fig. 3 shows a typical CNN architecture consisting of input, convolutional and subsampling layers, feature maps, fully connected layer and softmax regression (final stage). Convolutional and sub-sampling layers are arranged to reduce the computation time and to create spatial and configurable invariance gradually [21].

Convolutional layers consist of a number of filters. These filters convolute input from the previous layer with a set of weights and create an output called a feature map. The neurons in the filters are connected to the input data points and these points are multiplied with the weights. Because all neurons in the same filter share their weight, the optimization time and complexity of CNN is reduced [12].

$$C_{cn} = f(X * W_{cn} + b_{cn}) \tag{1}$$

Assuming that the convolutional layer input is $X \in R^{MxN}$, the layer output can be computed as in Eqs. 1. Here, M and N are dimensions of input data ; $*$ is convolution operater; C_{cn} is the cn-th feature map of the convolutional layer; X is the input data matrix; W_{cn} is the weight matrix of cn-th filter of the actual layer; b_{cn} is the cn-th bias; and f is nonlinear activation function that applied to the result [15].

To reduce the size of the features and parameters of the network by subsampling, the pooling layers come after the convolution lay-

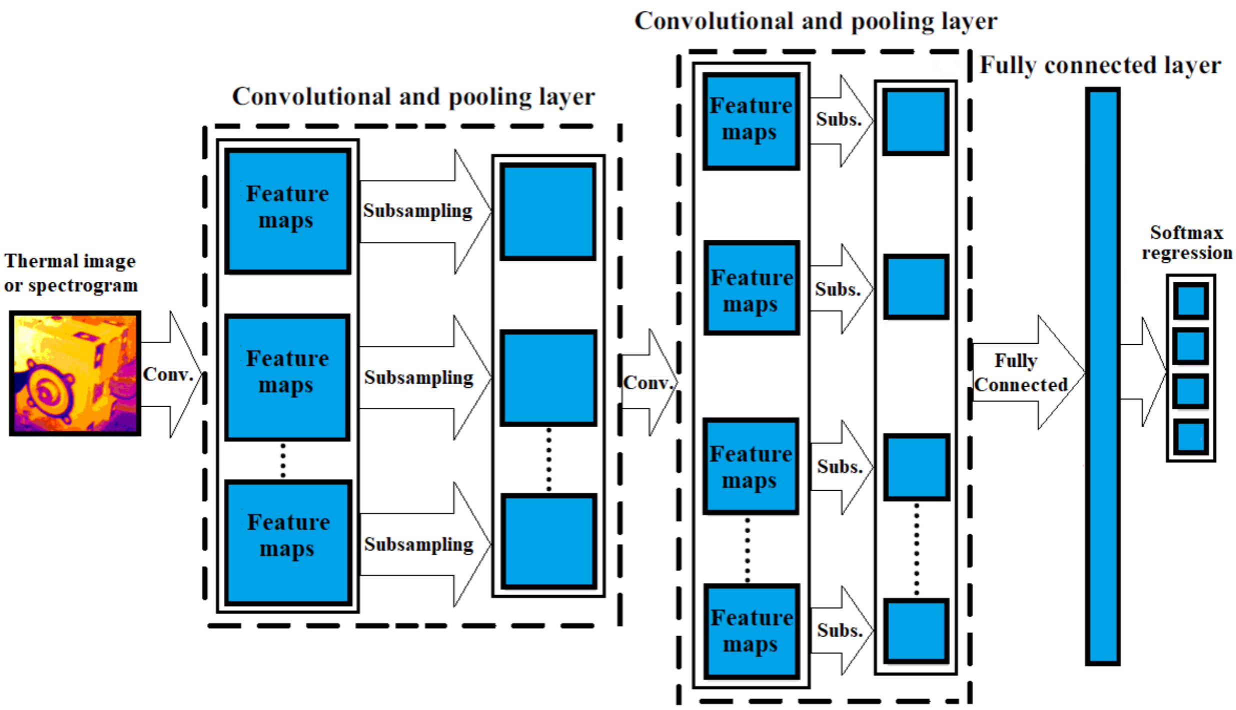


Fig. 3. Typical architecture of CNN

ers. There are three different pooling functions that calculate activation statistics: max pooling, mean pooling, and weighted pooling. Among these, max pooling is the most preferred function in CNN architecture [21]. S is pooling block size and $C_{cn} \in S$; P_{cn} is the output of the pooling layer; then max pooling activation can be written as in Eqs. 2 [15].

$$P_{cn} = \max C_{cn} \tag{2}$$

Finally, fully connected layer follows the combinations of convolutional and pooling layers. The fully connected layer is similar to a traditional neural network. So it can be applied to different classification problems. To achieve fast and accurate results, one hidden layer and softmax regression were chosen as the last layer. In this paper, as different gearboxes, which are healthy and faulty, are classified, the output of the softmax regression can be expressed as in Eqs. 3. Here, K is the label number, W_j is the weight matrix; b_j is bias; and O is the final result of the CNN [15].

$$O = \begin{bmatrix} P(y=1|x;W_1,b_1 \\ P(y=2|x;W_2,b_2 \\ \dots \\ P(y=K|x;W_K,b_K \end{bmatrix} = \frac{1}{\sum_{j=1}^{K} \exp(W_j x + b_j)} \begin{bmatrix} \exp(W_1 x + b_1) \\ \exp(W_2 x + b_2) \\ \dots \\ \exp(W_K x + b_K) \end{bmatrix} \tag{3}$$

3.2. Fault Diagnosis Protocol

When gears are in operation they are subject to dynamic operating loads that affect the temperature, vibration and sound magnitudes directly [1]. Therefore, the high temperature values or the increase in vibration or sound amplitudes of WGs cannot always be a cause of a fault in condition monitoring applications. Considering all working loads and speeds of the test rig [46], it should be demonstrated that the increase in temperature is due to fault or normal operating load (Table 1).

Table 1. Gearbox output speeds, loading rates and measurements

	WG Output Speeds	Loading Rates	Measurements
Healthy and faulty WG	30 rpm	0%, 25%, 50%, 75% and 100%	Thermal imaging, vibration measurement, and sound measurement
	50 rpm		
	70 rpm		

One healthy and three faulty WGs are selected. Faulty WGs were made by machining techniques. Healthy (H) and faulty wheels of gearboxes are given in Fig. 4. Faulty gearboxes are labelled as F1 (wear), F2 (pitting) and F3 (tooth breakage).

3.2.1. Infrared Thermography

Infrared thermography (IRT) has become one of the most effective condition monitoring tools [3], as they can enable reliable fault diagnosing results. Real-time and non-contact measurement of the temperature of machine equipment and elements can be carried out with IRT. In this way, it is possible to eliminate the failures occurring in the machines without causing catastrophic damage and production losses [30]. Under actual operating conditions, WGs are heated depending on the loading rate, operating speed and environment temperature. Measuring the radiant thermal energy distribution emitted from the surface of the WG and converting this energy distribution into a surface thermogram constitutes the basis of the condition monitoring with infrared thermography and gives information about the current operating status and possible gearbox faults such as wear, pitting, tooth breakage [23]. Therefore, thermal images of healthy and faulty WGs with 160x120 pixel were collected for the working conditions given in Table 1 and then, they were fed as visual data to train and validate CNN.

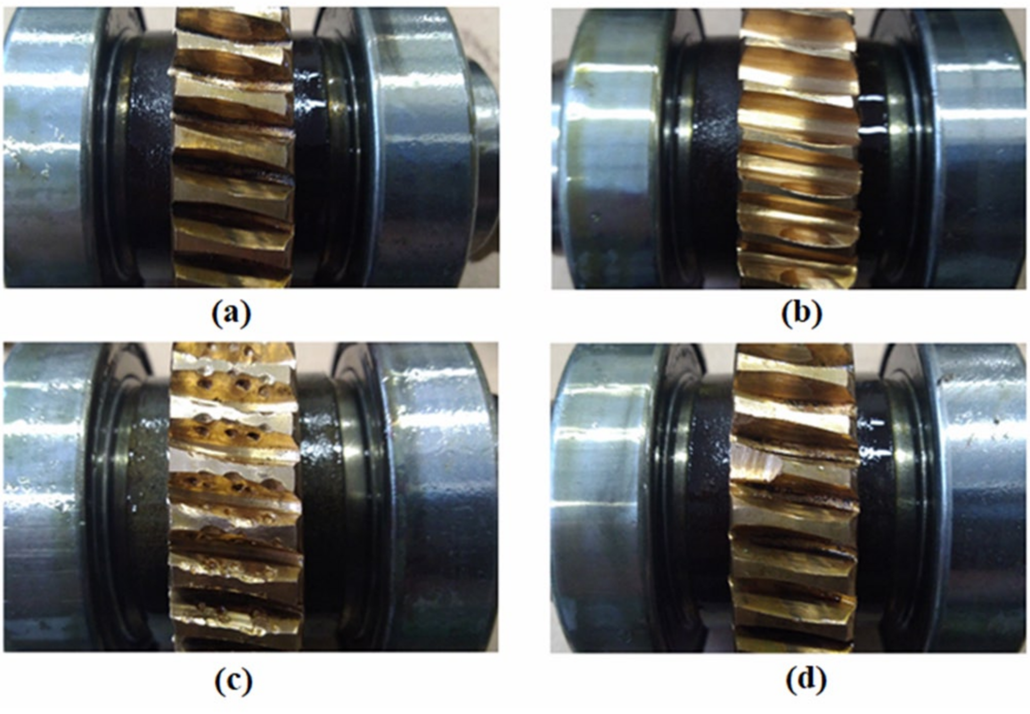


Fig. 4. Healthy (a), F1-wear (b), F2-pitting (c), F3-tooth breakage (d)

3.2.2. Short Time Fourier Transform and Spectrogram Images of Signals

Time-frequency analysis, one of the most common signal processing approach, can be used to understand the changes of sound and vibration signal components over time [6]. Fourier transform, is used to visualize the change of frequency components of the signal over time. This visual representation is called the spectrogram of the signal [38]. Vibration and sound spectograms were fed as visual data for CNN. In this study, short time Fourier transform (STFT) was preferred in order to obtain spectrograms.

In STFT, the signal, the function of time, is divided into short segments and the Fourier transform is applied for each segment. In the case of continuous time, the signal is multiplied by a window function when applying STFT operation as seen in Eqs. 4. Here, $x(t)$ is the time domain signal and $w(t)$ is the window function [44]:

$$STFT\{x(t)\}(\tau,\omega)=\int_{-\infty}^{+\infty}x(t)w(t-\tau)e^{-j\omega t}dt \tag{4}$$

STFT for discrete time is given as in Eqs. 5. Here, $x[n]$ is the signal and $w[n]$ is the window function. Based on this, the spectrogram of the STFT function (Fig. 5) is calculated as in Eqs. 6 [44]:

$$STFT\{x[t]\}(m,\omega)=X(m,\omega)=\sum_{-\infty}^{+\infty}x[n]w[n-m]e^{-j\omega n} \tag{5}$$

$$Spectrogram\{x[t]\}(m,\omega)=|X(m,\omega)|^2 \tag{6}$$

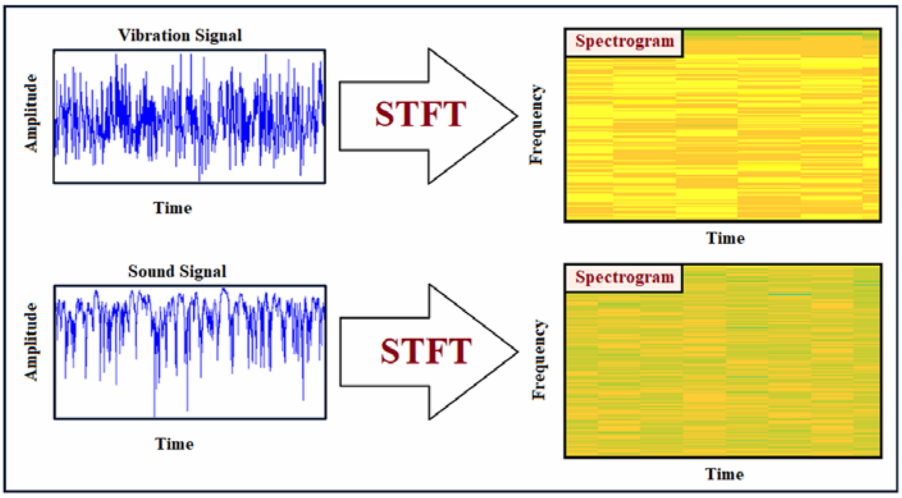


Fig. 5 Vibration and sound signals and spectrogram images

3.2.3. Fault Detection and Classification with Modified GoogLeNet Models

GoogLeNet architecture is used to construct CNN models in this study for the diagnosis of WG failures. GoogLeNet, offered by Szegedy et al. [42], is a deep convolutional neural network architecture based on new software technologies for classification and detection [42]. GoogLeNet is a pretrained CNN that has 22 layers deep, and the networks have an image input size of 224-by-224. The modified GoogLeNet properties and detailed structure of modified architecture are given in Table 2 and Table A1.

It is commonly agreed that the CNN mostly performs well with an enormous amount of data. However, it is possible to leverage deep learning even with limited data [4]. In some instances, where you can't gather more data due to cost or inconvenience, then there are some ways to follow during the training process. Fine tuning, data augmentation, using cosine loss function, or using autoencoders are some of them. This study has the data that can be considered as limited amount. So that, we did finetuning and data augmentation, respectively. In finetuning, we start with a pretrained model and updated all of the model's parameters for our new task, in essence retraining the whole model. All model parameters which are the number of classes in the dataset, the batch size used for training and the number of training epochs we want to run are updated. The train model function trains for the specified number of epochs and after each epoch runs a full validation step. After each epoch, the training and validation accuracies are reported. For data augmentation, we used a standard GoogleNet architecture that change the inputs in such a way that provides new data while not changing the label value. The 160x120 size raw input images are converted to 224x224x3 image input size with data augmentation.

Table 2. The modified GoogLeNet properties

Depth of layers	22
Number of layers	144
Number of connections	170
Input type	Image
Input size	224x224x3
Output type	Classification
Output size	4
Weight learn rate factor	10
Bias learn rate factor	10
Loss Function	Cross-entropy

Four different deep learning models were selected in this study (Table 3). They are thermal image based CNN(IRT-CNN), vibration signal based CNN(V-CNN), sound signal based CNN(S-CNN) and vibration-sound signals based CNN(VS-CNN). IRT-CNN is trained with raw thermal images; V-CNN is trained with vibration spectrogram images; S-CNN is trained with sound spectrogram images; and VS-CNN is trained with both of vibration and sound spectrogram images. The technical parameters; validation frequency is selected as 6 Hz and the learning rate is 0.001 for all models in the table.

A simple flowchart of training process and the images used for the classification of faults with different CNN models are given in Fig. 6. Accordingly, thermal images and signal spectrogram images of H, F1, F2 and F3 gearboxes are collected separately for all operating conditions at first. The data of each gearbox is divided into three groups as train, validation and test. GoogLeNet outputs are modified to classify healthy and faulty gearboxes. Using train and validation data, CNN

Table 3. Experimental parameters for different deep learning models

<table>
<tr><th>IRT-CNN</th><th>V-CNN</th><th>S-CNN</th><th>VS-CNN</th><th colspan="2">Fault Diagnosis</th></tr>
<tr><th colspan="4">Input Parameters</th><th colspan="2">Output Parameters</th></tr>
<tr><td rowspan="2">Raw thermal images</td><td rowspan="2">Vibration spectrogram images</td><td rowspan="2">Sound spectrogram images</td><td rowspan="2">Vibration-sound spectrogram images</td><td colspan="2">Healthy (H)</td></tr>
<tr><td rowspan="5">Fault Detection</td><td>Fault 1 (F1)</td></tr>
<tr><td colspan="4">Train, Validation and Test Samples</td><td>Fault 2 (F2)</td></tr>
<tr><td colspan="4">50%; 25%; 25%</td><td rowspan="3">Fault 3 (F3)</td></tr>
<tr><td colspan="2">Loading Rates (LR)</td><td colspan="2">0%; 25%; 50%; 75%; 100%</td></tr>
<tr><td colspan="2">Gearbox output speed (GOS)</td><td colspan="2">30 rpm; 50 rpm; 70 rpm</td></tr>
</table>

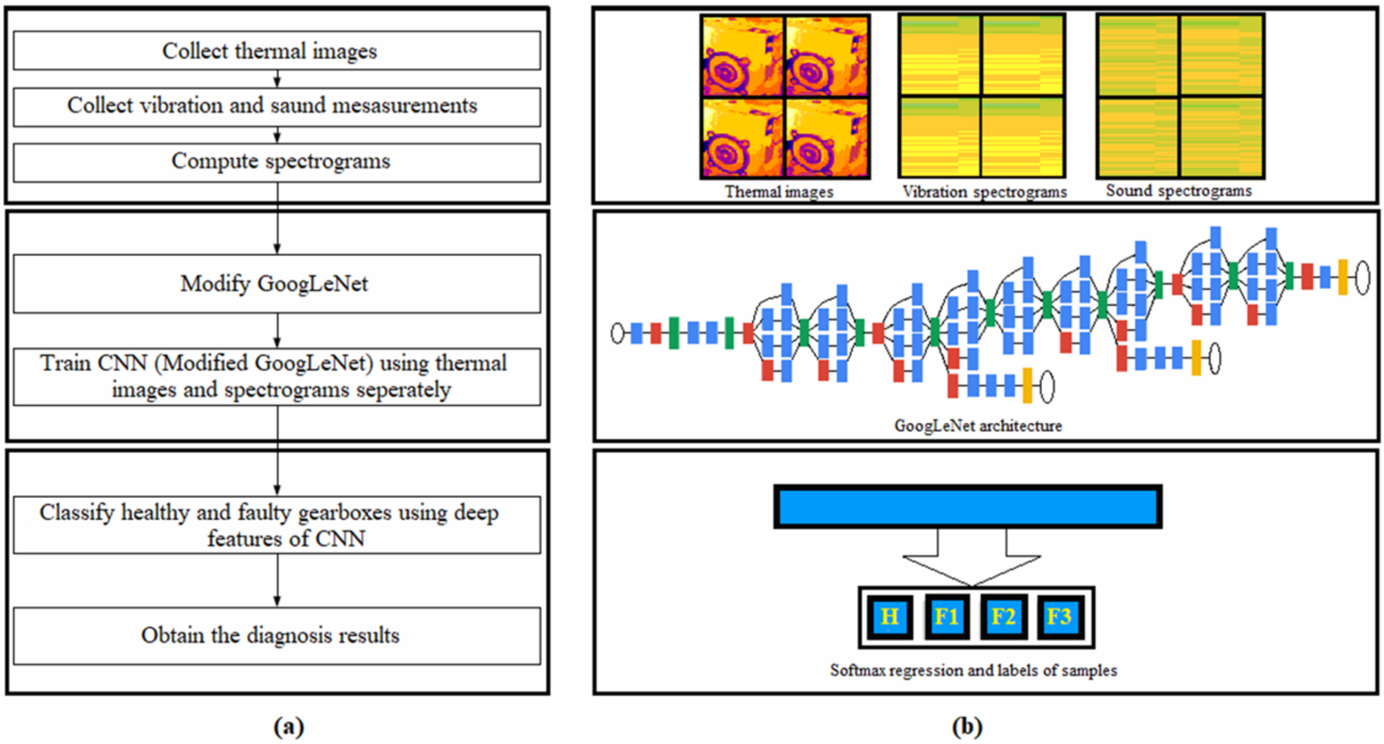


Fig. 6. CNN-based fault diagnosis of WG (a) flowchart of training process, and (b) Images

is trained and classification is performed based on deep features. Finally, the trained network is tested and fault diagnosis is carried out according to data labels.

4. Result and Discussions

This section contains the findings of the experiments and a brief discussion on the analysis based on past and current studies.

4.1. Findings of Thermal Imaging and Time-Frequency Signal Analysis

A total of 120 thermal images were collected in different loading rates (LR) and gearbox output speeds (GOS) for each gearbox. A total of 480 thermal images were obtained from all gearboxes for training, validation and testing of the IRT-CNN model.

The difference between the temperature distributions on the surfaces of healthy and faulty gearboxes increases gradually with increasing LR and GOS. For all LR and GOS, the temperature distributions of F1, F2 and F3 gearboxes are higher than H. The thermal images of H, F1, F2 and F3 for GOS = 50 rpm and LR = 50% are given in Fig. 7. The maximum temperature for F1 is 105°C and the average surface temperature is 75°C. The hottest gearbox is observed as F1 during the experiments. The maximum temperature value for F2 is 93°C and the average surface temperature is 66°C. F2 is hotter than F3 for the same conditions. The maximum temperature value for F3 is 88°C and the average surface temperature is 64°C. The less heated gearbox is observed as H. The maximum temperature value for H is 83°C and the average surface temperature is 60 °C. For F1, F2, F3 and H gearboxes, the minimum surface temperature values measured under these conditions are 18.4°C; 16.7° C; 16.1°C; 15.7 °C, respectively. Accordingly, the worn gearbox is the hottest gearbox. Pitting failure causes more heat than the broken tooth. The healthy gearbox is the coolest one during experiments.

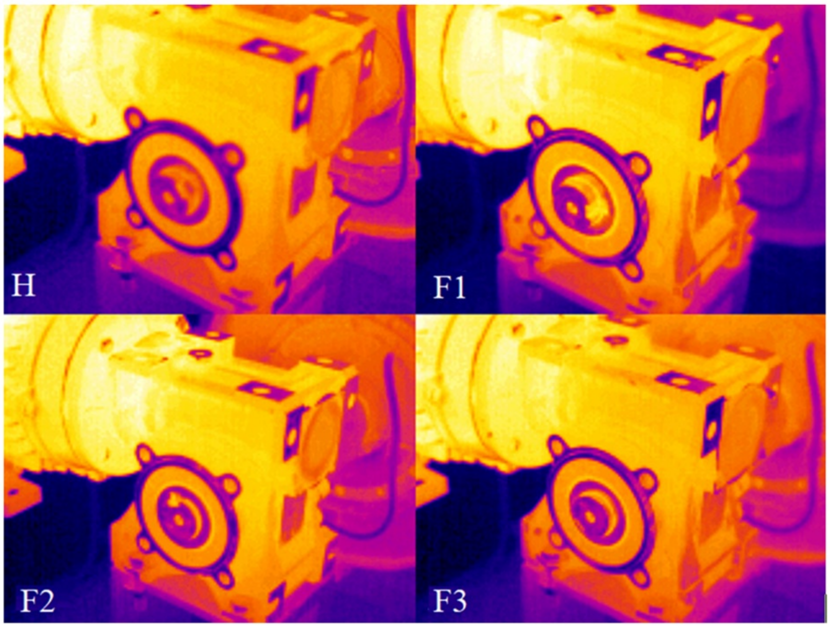


Fig. 7. The thermal images of H, F1, F2 and F3 gearboxes

The number of vibration spectrogram images were 120 for each gearbox in different LR and GOS. A total of 480 vibration spectrograms were obtained from all gearboxes for training, validation and testing of the V-CNN model. In Fig. 8, time waveforms and spectrogram images for H, F1, F2 and F3 gearboxes are given in the interval of 0-1 s for LR = 50%, GOS = 50 rpm. The sampling rate for time-frequency analysis was chosen as 1600 Hz and frequencies higher than 1000 Hz were eliminated by Butterworth low-pass filter for vibration and sound analysis. Fig. 9, shows the time waveforms and spectrogram images of sound measurements for LR = 50%, GOS = 50 rpm. The time interval for sound waveform is 0-0.4 s. A total of 480 sound spectrogram images were also collected for training, validation and testing of the S-CNN model. It was seen that the amplitude of sound and vibration signals were increasing with increasing load and speed. Moreover, different faulty gearboxes produced different signal amplitudes as in Fig. 8 and Fig. 9 that the amplitudes of the sound and vibration signal of the faulty gearboxes are higher than the healthy gearbox. From the figures, F1 has the highest vibration amplitudes. These amplitudes vary from −31.7 m/s^2 to 42.6 m/s^2 in the 0-1 s time interval and 0-750 Hz frequency range. The H gearbox vibration signal amplitudes range from −28.4 m/s^2 to −29.4 m/s^2 and they are the lowest in the same conditions. The vibration amplitudes of the F2 generally vary from −32.4 m/s^2 to 30 m/s^2 and larger than amplitudes of the F3. In addition, F3 has the highest sound amplitudes, ranging from −13 dB to 75 dB in the 0-0.4 s time interval and 0-1900 Hz frequency range. H reducer has the lowest sound amplitudes varying from −10.1 dB to 68 dB in the same conditions. The sound amplitudes of the F2 are generally between −9 dB and 72 dB which are larger than the amplitudes of the F3.

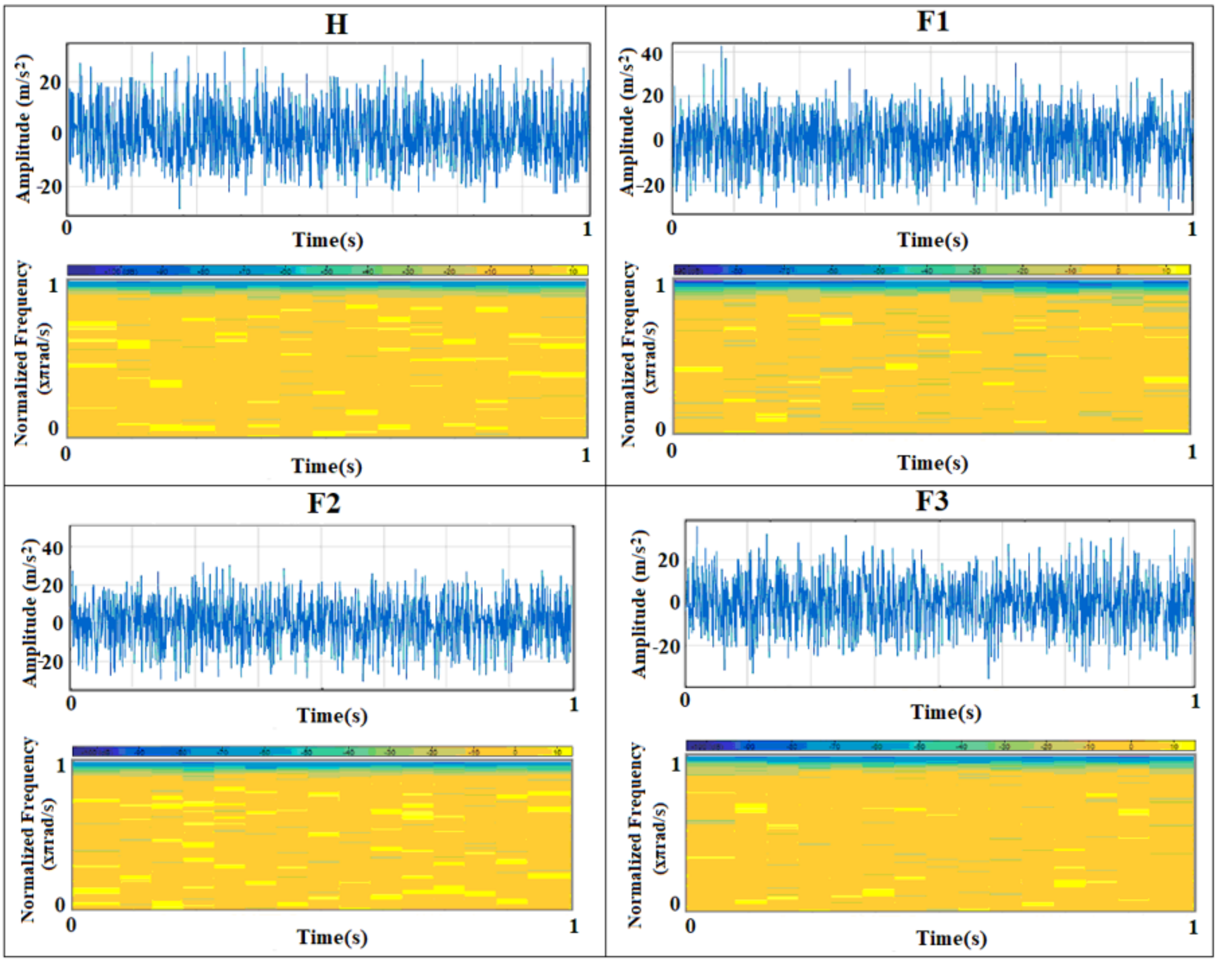


Fig. 8. Time waveforms and spectrogram images of vibration measurements

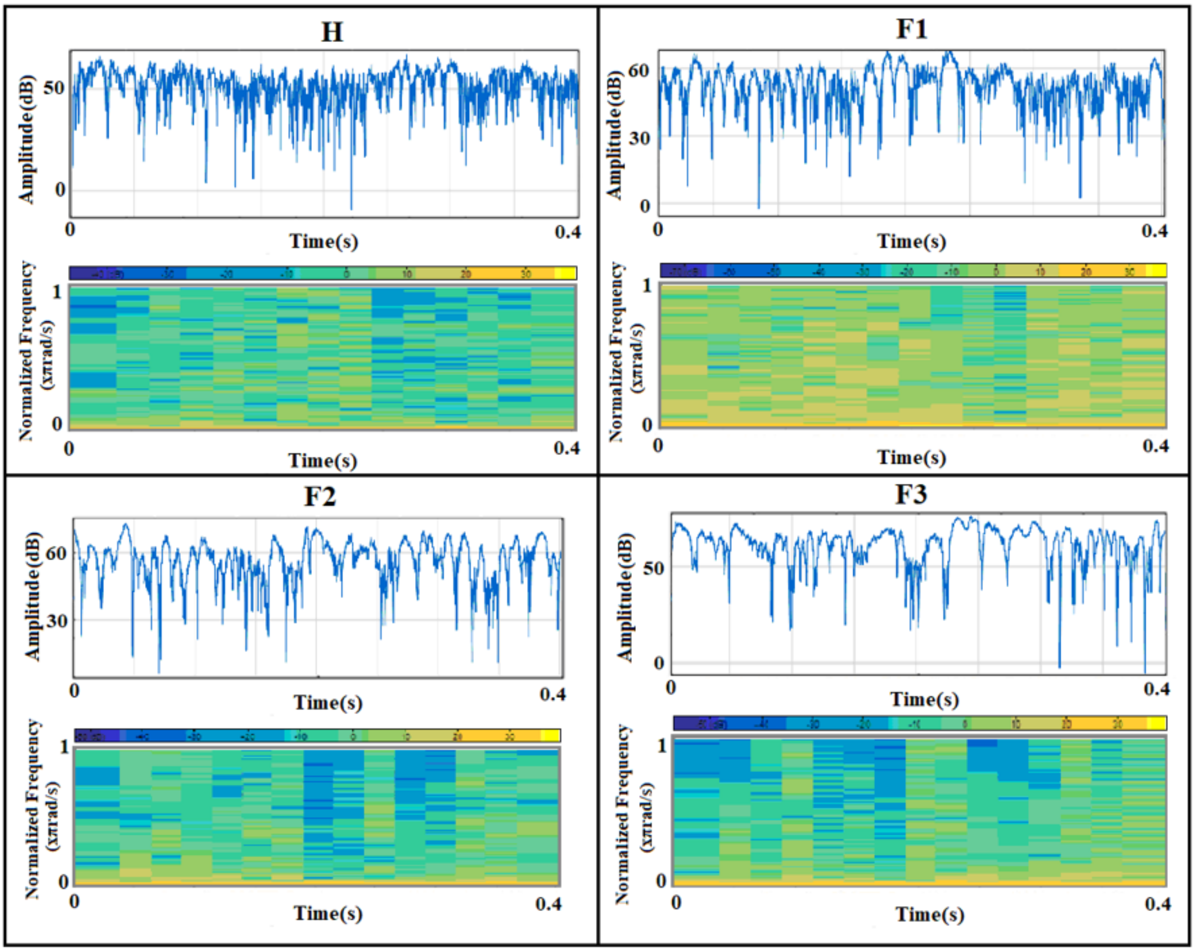


Fig. 9. Time waveforms and spectrogram images of sound measurements

4.2. Results of Fault Diagnosis with Different CNN Models

The spectrogram images of vibration and sound signals were used together in training, validation and test of the VS-CNN model. The targets for CNN models are healthy and faulty gearboxes. Table 4 shows class labels of gearboxes and the number of samples used for training, validation and testing of different CNN models.

Fig. 10 shows the accuracy rates for different CNN models. Accordingly, the highest training and validation accuracy rates were achieved with the IRT-CNN model based on thermal images. It reached 100% in the 30th iteration and 10 epochs. This success rate is 87.5% for V-CNN and 81.67% for S-CNN model. VS-CNN, which uses spectrogram images of vibration and sound samples together, has the lowest training and validation accuracy with 73.33%. Li et al. [24] achieved similar results in their study for the condition monitoring of bevel gearbox. However, they did not use spectrograms for training and validation of CNN models. In addition, their experiments were carried out under constant speed and loading rate, and the effect of real working conditions on temperature change was not taken into account sufficiently (Table 6, Table 7).

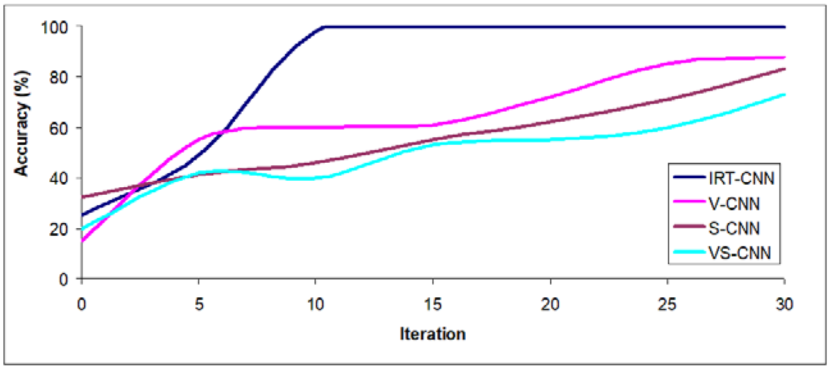


Fig. 10. Accuracy rates for different CNN models

Table 4. Class labels and number of samples for CNN models

Class Label	Number of Training Samples	Number of Validation Samples	Number of Test Samples
H	60	30	30
F1	60	30	30
F2	60	30	30
F3	60	30	30

Fig. 11 shows the losses in training and validation progresses for different CNN models. Losses are a quantitative measure of the difference between the predicted output and the actual output,

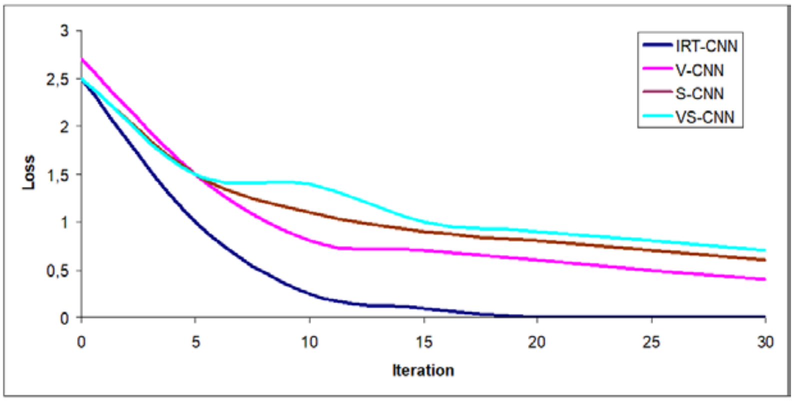


Fig. 11. Losses for different CNN models

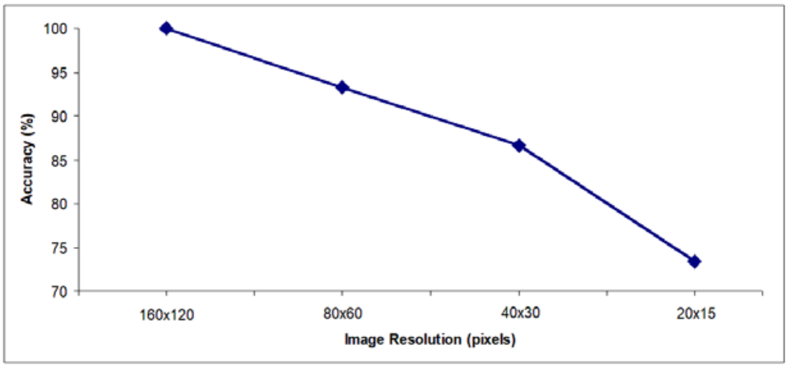


Fig. 12. The change of accuracy rates due to thermal image resolutions

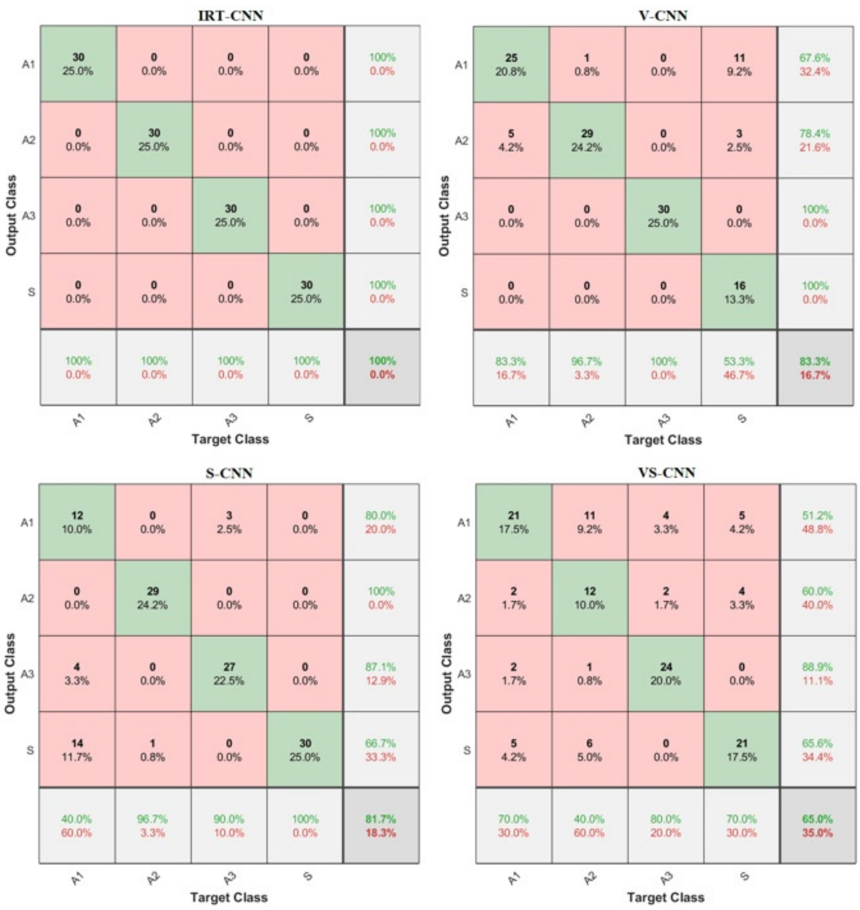


Fig. 13. Confusion matrices of CNN models

and decrease with the number of iterations. Depending on the suitability and adequacy of the data set, losses and the number of iterations decrease. Accordingly, the highest loss occurred during the training progress of the VS-CNN model. Training and verification loss for the IRT-CNN model decreased to 0% in the 30th iteration and 10 epochs.

To test CNN models, 25% of the samples collected from the gearboxes was used. Confusion matrices of different CNN models are given in Fig. 13. The IRT-CNN model classified all test inputs for H, F1, F2 and F3 gearboxes correctly. For this success rate, the test was repeated 10 times and the standard deviation was found to be 0. Whereas the confusion matrix of V-CNN model's success rate is 83.3% and half of the samples of the H gearbox is estimated incorrectly. When the S-CNN model is tested, 81.7% of the samples are estimated correctly. The rate of misclassification of F1 is high. The lowest test success rate belongs to VS-CNN model. In this model, only 65% of all test samples were estimated correctly. It can be inferred from confusion matrix that the test samples of the A2 gearbox are classified incorrect considerably.

Table 5 shows the success rates of CNN models trained and tested with different numbers of samples. Accordingly, as the number of samples increases, validation and test accuracy rates and training time of all CNN models increase. The highest validation and test accuracy rates for all models are achieved when 120 samples are used. Validation and test accuracy rates of IRT-CNN models are higher than all V-CNN, S-CNN and VS-CNN models. It is notable that even if the IRT-CNN training sample numbers are 30 or 60, it still has a very high diagnostic success (89-90%). %).The accuracy rates with the increasing amount of training and testing data of V-CNN and S-CNN models are more affected compared to IRT-CNN models. Even with the least number of samples, IRT-CNN has 89.3 % accuracy.

One other topic searched in this study is the resolution of images that are to be classified. Resolution of the image varies due to different input sources, different imaging devices and different environment noises [16] . Variation in images resolution alters the visual information of images [7]. Fig. 12 shows the variation of accuracy rates depending on the image resolutions. When the performance comparison of classifier for IRT-CNN dataset analysed from Figure 12, it is noticeable that accuracy rate decreases when image resolution decreases. The results are in accordance with the literature [7].

These results have shown that the IRT-CNN algorithm is effective for using the original temperature values for different fault diagnosis. It is worth noting that this study is conducted as a special case study in a factory where the unwanted working noises are surpassed. However, the system was tried to be made as a real as possible. Further effort can be put and the method can be tried on a real shift.

Table 6 and Table 7 show a general comparison of CNN models proposed by Li et al. [24] and CNN models proposed from this study.

Table 5. Effect of the number of samples on the performance of CNN models

CNN Models	Number of Samples	Validation Accuracy (%)	Test Accuracy (%)	Training Time (s)
IRT-CNN	30	90.63	89.3	122
IRT-CNN	60	95	90	123
IRT-CNN	120	100	100	366
V-CNN	30	59.38	57.1	129
V-CNN	60	65	61.7	145
V-CNN	120	87.50	83.3	456
S-CNN	30	59.38	42.9	127
S-CNN	60	58.33	41.7	142
S-CNN	120	81.67	81.7	395
VS-CNN	30	40.63	39.3	128
VS-CNN	60	46.67	55	147
VS-CNN	120	73.33	65	423

Table 6. The proposed CNN models in the study

CNN Models	Test Accuracy (%)	Training Time (s)
IRT-CNN	100	366
V-CNN	83.3	456
S-CNN	81.7	395
VS-CNN	65	423

Table 7. CNN models proposed by Li et al. [24]

CNN Models	Test Accuracy (%)	Training Time (s)
IRT-CNN	100	470
V-CNN	71.53	542

5. Conclusions

In this paper, a CNN based condition monitoring study based on infrared thermal images, vibration spectrogram images and sound spectrogram images was performed and the model is used to diagnose different types of faults in WGs. In real working conditions with a test rig, which can operate at different speed and loading rates, the CNN model based on thermal imaging yielded the most successful results than other models. Thermal images collected at different loading rates and different speeds increased the success rate of IRT-CNN. The operating conditions directly affect the temperature of the gearbox and the thermal image pattern. Also, different types of faults tend to show different temperature distribution under different operating conditions. The results showed that healthy and faulty gearbox can be correctly classified 100% with the IRT-CNN model; 83.3% with the V-CNN model; 81.7% with the S-CNN model; and 65% with the VS-CNN model. Based on this comparison, it was found that the use of thermal images with convolutional neural networks (CNN) generates the highest classification rates. With the high resolution detector thermal cameras, successful fault detection can be made with small data sets considering different operating conditions. A new possible advantage of the IRT-CNN model is that it allows remote fault diagnosis. It can be useful to work with the thermal cameras, which are relatively cheaper than other signal processing devices. Faults of WG can be accurately detected with loss of cost, time and money.

References

1. Al-Arbi S K. Condition Monitoring of Gear Systems using Vibration Analysis, Doctoral Dissertation. University of Huddersfield, 2012.
2. Al-Musawi A K, Anayi F, Packianather M. Three-Phase Induction Motor Fault detection based on Thermal Image Segmentation. Infrared Physics & Technology 2019; 104: 103140, https://doi.org/10.1016/j.infrared.2019.103140.
3. Bagavathiappan S, Lahiri B B, Saravanan T, Philip J, Jayakumar T. Infrared thermography for condition monitoring, A review. Infrared Physics & Technology 2013; 60: 35-55, https://doi.org/10.1016/j.infrared.2013.03.006.
4. Barz B, Denzler J. Deep learning on small datasets without pre-training using cosine loss. The IEEE Winter Conference on Applications of Computer Vision 2020; 1371-1380, https://doi.org/10.1109/WACV45572.2020.9093286.
5. Carden E P, Fanning P. Vibration based condition monitoring: a review. Structural health monitoring 2004; 3(4): 355-377, https://doi.org/10.1177/1475921704047500.
6. Cohen L. Time-Frequency Analysis, New Jersey: Prentice Hall PTR, 1995.
7. Dodge S, Karam L. Understanding how image quality affects deep neural networks. In 2016 eighth international conference on quality of multimedia experience (QoMEX) IEEE 2016; 1-6, https://doi.org/10.1109/QoMEX.2016.7498955.
8. Ghodake S, Mishra A K, Deokar A V. A Review Paper on Fault Detection of Worm Gearbox. International Advanced Research Journal in Science, Engineering and Technology 2016; 3(1): 161-164.
9. Goyal D, Pabla B S, Dhami S S. Condition monitoring parameters for fault diagnosis of fixed axis gearbox: a review. Archives of Computational Methods in Engineering 2017; 24(3): 543-556, https://doi.org/10.1007/s11831-016-9176-1.
10. Goyal D, Pabla B S, Dhami S S. Non-contact sensor placement strategy for condition monitoring of rotating machine-elements. Engineering Science and Technology 2019; 22(2): 489-501, https://doi.org/10.1016/j.jestch.2018.12.006.
11. Haberhauer H, Bodenstein F. Maschinenelemente. Berlin: Springer, 2009.

12. Hoang D T, Kang H J. Rolling element bearing fault diagnosis using convolutional neural network and vibration image. Cognitive Systems Research 2019; 53: 42-50, https://doi.org/10.1016/j.cogsys.2018.03.002.
13. Janssens O, Schulz R, Slavkovikj V, Stockman K, Loccufier M, Van de Walle R, Van Hoecke S. Thermal image based fault diagnosis for rotating machinery. Infrared Physics & Technology 2015; 73: 78-87, https://doi.org/10.1016/j.infrared.2015.09.004.
14. Janssens O, Van de Walle R, Loccufier M, Van Hoecke S. Deep learning for infrared thermal image based machine health monitoring. IEEE/ASME Transactions on Mechatronics 2017; 23(1): 151-159, https://doi.org/10.1109/TMECH.2017.2722479.
15. Jing L, Zhao M, Li P, Xu X. A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. Measurement 2017; 111: 1-10, https://doi.org/10.1016/j.measurement.2017.07.017.
16. Kannojia S P, Jaiswal G. Effects of Varying Resolution on Performance of CNN based Image Classification: An Experimental Study. International Journal of Computer Sciences and Engineering 2018; 6(9): 451-456, https://doi.org/10.26438/ijcse/v6i9.451456.
17. Kłosowski G, Rymar czyk T, Kania K, Świć A, Cieplak T. Maintenance of industrial reactors supported by deep learning driven ultrasound tomography. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(1): 138-147, https://doi.org/10.17531/ein.2020.1.16.
18. Lei Y, Lin J, He Z, Zuo M J. A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mechanical Systems and Signal Processing 2013; 35(1-2): 108-126, https://doi.org/10.1016/j.ymssp.2012.09.015.
19. Li C, Sanchez R V, Zurita G, Cerrada M, Cabrera D, Vásquez R E. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. Mechanical Systems and Signal Processing 2016; 76: 283-293, https://doi.org/10.1016/j.ymssp.2016.02.007.
20. Li C, Sánchez R V, Zurita G, Cerrada M, Cabrera D. Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning. Sensors 2016; 16(6): 895, https://doi.org/10.3390/s16060895.
21. Li C, Sanchez R V. Gearbox fault identification and classification with convolutional neural Networks. Shock and Vibration 2015; Article ID: 390134, https://doi.org/10.1155/2015/390134.
22. Li X, Li J, He D, Qu Y. Gear pitting fault diagnosis using raw acoustic emission signal based on deep learning. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(3): 403-410, https://doi.org/10.17531/ein.2019.3.6.
23. Li Y, Du X, Wan F, Wang X, Yu H. Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging. Chinese Journal of Aeronautics 2020; 33(2): 427-438, https://doi.org/10.1016/j.cja.2019.08.014.
24. Li Y, Gu J X, Zhen D, Xu M, Ball A. An Evaluation of Gearbox Condition Monitoring Using Infrared Thermal Images Applied with Convolutional Neural Networks. Sensors 2019; 19(9): 2205, https://doi.org/10.3390/s19092205.
25. Li Y, Wang K. Modified convolutional neural network with global average pooling for intelligent fault diagnosis of industrial gearbox. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(1): 63-72, https://doi.org/10.17531/ein.2020.1.8.
26. Liu Z, Zhang L. A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings. Measurement 2020; 149: 107002, https://doi.org/10.1016/j.measurement.2019.107002.
27. Loutas T H, Roulias D, Pauly E, Kostopoulos V. The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery. Mechanical Systems and Signal Processing 2011; 25(4): 1339-1352, https://doi.org/10.1016/j.ymssp.2010.11.007.
28. Loutas T H, Sotiriades G, Kalaitzoglou I, Kostopoulos V. Condition monitoring of a single-stage gearbox with artificially induced gear cracks utilizing on-line vibration and acoustic emission measurements. Applied Acoustics 2009; 70(9): 1148-1159, https://doi.org/10.1016/j.apacoust.2009.04.007.
29. Maldague X. Theory and Practice of Infrared Technology for Non-Destructive Testing. New York: John Wiley-Interscience, 2001.
30. Meola C. Infrared thermography recent advances and future trends. Sharjah: Bentham Science Publishers, 2012, https://doi.org/10.2174/97816080514341120101.
31. Mohanty A R. Machinery Condition Monitoring: Principles and Practices. New York: CRC Press, Taylor&Francis Group, 2015, https://doi.org/10.1201/9781351228626.
32. Osman S, Wang W. An enhanced Hilbert-Huang transform technique for bearing condition monitoring. Measurement Science and Technology 2013; 24(8): 085004, https://doi.org/10.1088/0957-0233/24/8/085004.
33. Randall R B. Vibration-based Condition Monitoring: Industrial, Aerospace and Automotive Applications. West Sussex: John Wiley & Sons Ltd., 2011, https://doi.org/10.1002/9780470977668.
34. Ring E F J, Ammer K. Infrared thermal imaging in medicine. Physiological Measurement 2012; 33(3): R33, https://doi.org/10.1088/0967-3334/33/3/R33.
35. Sait A S, Sharaf-Eldeen Y I. A review of gearbox condition monitoring based on vibration analysis techniques diagnostics and prognostics. In: Proulx T (Ed.), Rotating Machinery, Structural Health Monitoring, Shock and Vibration 2011; 5: 307-324, https://doi.org/10.1007/978-1-4419-9428-8_25.
36. Sasai S, Zhen Y X, Suetake T, Tanita Y, Omata S, Tagami H. Palpation of the skin with a robot finger: an attempt to measure skin stiffness with a probe loaded with a newly developed tactile vibration sensor and displacement sensor. Skin Research and Technology 1999; 5(4): 237-246, https://doi.org/10.1111/j.1600-0846.1999.tb00136.x.
37. Sharma V, Parey A. A review of gear fault diagnosis using various condition indicators. Procedia Engineering 2016; 144: 253-263, https://doi.org/10.1016/j.proeng.2016.05.131.
38. Sharma V, Parey A. Frequency domain averaging based experimental evaluation of gear fault without tachometer for fluctuating speed conditions. Mechanical Systems and Signal Processing 2017; 85: 278-295, https://doi.org/10.1016/j.ymssp.2016.08.015.
39. Singh G, Kumar T C A, Naikan V N A. Induction motor inter turn fault detection using infrared thermographic analysis. Infrared Physics & Technology 2016; 77: 277-282, https://doi.org/10.1016/j.infrared.2016.06.010.
40. Singh G, Naikan V N A. Infrared thermography based diagnosis of inter-turn fault and cooling system failure in three phase induction motor. Infrared Physics & Technology 2017; 87: 134-138, https://doi.org/10.1016/j.infrared.2017.10.007.
41. Sun B, Li M M, Liao B P, Yang X, Cao Y T, Cui B F, Feng Q, Ren Y, Yang, D Z. Time-variant reliability modeling based on hybrid non-probability method. Archive of Applied Mechanics 2020; 90(2): 209-219, https://doi.org/10.1007/s00419-019-01605-1.
42. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: 28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston MA, 2015: 1-9, https://doi.org/10.1109/CVPR.2015.7298594.

43. Tanaka Y, Horita Y, Sano A. Finger-mounted skin vibration sensor for active touch. In: Isokoski P, Springare J. (Eds.), EuroHaptics: International Conference on Human Haptic Sensing and Touch Enabled Computer Applications. Tampere, Finland, 2012: 169-174, https://doi.org/10.1007/978-3-642-31404-9_29.
44. Vashisht R K, Peng Q. Crack detection in the rotor ball bearing system using switching control strategy and Short Time Fourier Transform. Journal of Sound and Vibration 2018; 432: 502-529, https://doi.org/10.1016/j.jsv.2018.06.061.
45. Wang J, Gao R X, Yan R. Multi-scale enveloping order spectrogram for rotating machine health diagnosis. Mechanical Systems and Signal Processing 2014; 46(1): 28-44, https://doi.org/10.1016/j.ymssp.2013.06.001.
46. Waqar T, Demetgul M. Thermal analysis MLP neural network based fault diagnosis on worm gears. Measurement 2016; 86: 56-66, https://doi.org/10.1016/j.measurement.2016.02.024.
47. Younus A M, Yang B S, Intelligent fault diagnosis of rotating machinery using infrared thermal image. Expert Systems with Applications 2012; 39(2): 2082-2091, https://doi.org/10.1016/j.eswa.2011.08.004.
48. Zhang X, Kang J, Bechhoefer E, Zhao J. A new feature extraction method for gear fault diagnosis and prognosis. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16(2): 295-300.
49. Zhang X, Zhao J. Compound fault detection in gearbox based on time synchronous resample and adaptive variational mode decomposition. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(1): 161-169, https://doi.org/10.17531/ein.2020.1.19.
50. Zhang Q, Zhao W, Xiao S G. Fault Diagnosis of Gear Based on Singular Value Decomposition and RBF Neural Network. 2nd International Conference on Frontiers of Sensors Technologies 2017; 470-474, https://doi.org/10.1109/ICFST.2017.8210559.
51. Zhao R, Wang D Z, Yan R Q, Mao K Z, Shen F, Wang J J. Machine Health Monitoring Using Local Feature-Based Gated Recurrent Unit Networks. IEEE Transactions on Industrial Electronics 2018; 65: 1539-1548, https://doi.org/10.1109/TIE.2017.2733438.

Appendix A

Table A1. Detailed structure of modified GoogLeNet

Layer type	Patch size/ stride	Output size	Depth	#1x1	#3x3 reduce	#3x3	#5x5 reduce	#5x5	Pool proj.
Convolution	7x7/2	112x112x64	1	-	-	-	-	-	-
Max pool	3x3/2	56x56x64	0	-	-	-	-	-	-
Convolution	3x3/1	56x56x192	2	-	64	192	-	-	-
Max pool	3x3/2	28x28x192	0	-	-	-	-	-	-
Inception(3a)	-	28x28x256	2	64	96	128	16	32	32
Inception(3b)	-	28x28x480	2	128	128	192	32	96	64
Max pool	3x3/2	14x14x480	0	-	-	-	-	-	-
Inception(4a)	-	14x14x512	2	192	96	208	16	48	64
Inception(4b)	-	14x14x512	2	160	112	224	24	64	64
Inception(4c)	-	14x14x512	2	128	128	256	24	64	64
Inception(4d)	-	14x14x528	2	112	144	288	32	64	64
Inception(4e)	-	14x14x832	2	256	160	320	32	128	128
Max pool	3x3/2	7x7x832	0	-	-	-	-	-	-
Inception(5a)	-	7x7x832	2	256	160	320	32	128	128
Inception(5b)	-	7x7x1024	2	384	192	384	48	128	128
Avg pool	7x7/1	1x1x1024	0	-	-	-	-	-	-
Dropout (40%)	-	1x1x1024	0	-	-	-	-	-	-
Fully connected	-	1x1x4	1	-	-	-	-	-	-
Softmax	-	1x1x4	0	-	-	-	-	-	-

Yunus Emre KARABACAK
General Directorate of Tea Enterprises
53080 Rize, Turkey

Nurhan GÜRSEL ÖZMEN
Levent GÜMÜŞEL
Karadeniz Technical University, Mechanical Engineering Department
61080 Trabzon, Turkey

E-mails: Y_emre_karabacak@hotmail.com, gnurhan@ktu.edu.tr, gumusel@ktu.edu.tr

Article citation info:
UŁANOWICZ L, JASTRZĘBSKI G, SZCZEPANIAK P. Method for estimating the durability of aviation hydraulic drives. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 557–564, http://dx.doi.org/10.17531/ein.2020.3.19.

Leszek UŁANOWICZ
Grzegorz JASTRZĘBSKI
Paweł SZCZEPANIAK

METHOD FOR ESTIMATING THE DURABILITY OF AVIATION HYDRAULIC DRIVES

METODA SZACOWANIA TRWAŁOŚCI LOTNICZYCH NAPĘDÓW HYDRAULICZNYCH*

Throughout previous practice, estimating the life of aviation hydraulic drive assemblies has been utilizing a variant, which requires conducting long-lasting studies of the drive assemblies until they move to the unfitness state. Such studies, which enable estimating life a posteriori, are costly and long-lasting. Hence the need to look for new strategies for estimating life. The article presents a method of estimating the durability of a hydraulic drive assembly based on the control of its change in technical condition. Inspection of the technical condition enables timely detection of the condition before the emergency hydraulic assembly. The novelty of the method is to use, to detect the condition before the emergency team, the principle of determining the pre-emptive control parameter tolerance. Pre-emptive tolerances are a set of control parameter values between threshold levels and pre-emergency (allowable) levels. The intensity of depletion of durability (intensity of aging, wear) is random. The paper presents a stochastic description of the control parameter change and the resulting empirical relationships between the control parameter verification time probability density (verification periodicity) and the control parameter value change probability density. The inter-relations between these two functions were described. It also presents empirical relationships enabling the determination of the permissible value for the control parameters and the periodicity of the control parameter checks after exceeding the limit value. An example of estimating the life of a hydraulic piston pump on-board an aircraft operated in the Polish Air Forces was shown. The permissible values and the time for the first control parameter verification after exceeding the limit value were determined for selected control parameters of the hydraulic pump. The proposed method binds life (fitness time) with the physical wear mechanisms concerning the assemblies. It can be applied in work aimed at determining the resource life of technical equipment. Furthermore, it enables utilizing technical equipment according to a technical state strategy with monitoring the parameters.

Keywords: *aviation, lifetime, hydraulic drive, hydraulic pump, technical condition.*

W dotychczasowej praktyce szacowania trwałości zespołów lotniczych napędów hydraulicznych stosowany jest wariant, który wymaga prowadzenia długotrwałych badań zespołów napędu do czasu ich przejścia w stan niezdatności. Badania tego typu, umożliwiające szacowanie trwałości a posteriori, są kosztowne i długotrwałe. Istnieje więc potrzeba poszukiwania nowych strategii szacowania trwałości. W artykule zaprezentowano metodę szacowania trwałości zespołu napędu hydraulicznego opartą o kontrolę jego zmiany stanu technicznego. Kontrola stanu technicznego umożliwia wykrycie we właściwym czasie stanu przed awaryjnego zespołu hydraulicznego. Novum metody jest wykorzystanie, do wykrycie stanu przed awaryjnego zespołu, zasady wyznaczania uprzedzających tolerancji parametru kontrolnego. Tolerancje uprzedzające stanowią zbiór wartości parametru kontrolnego zawartych między poziomami granicznym i przed awaryjnym (dopuszczalnym). Intensywność wyczerpywania się trwałości (intensywności starzenia, zużywania) ma losowy charakter. W artykule przedstawiono stochastyczny opis zmiany parametru kontrolnego oraz wynikające z niego empiryczne zależności funkcji gęstości prawdopodobieństwa czasu przeprowadzania sprawdzeń parametru kontrolnego (okresowość kontroli) i funkcji gęstości prawdopodobieństwa zmiany wartości parametru kontrolnego. Opisano wzajemne związki obu tych funkcji. Przedstawiono zależności umożliwiające wyznaczenie wartości dopuszczalnej parametru kontrolnego i okresowość sprawdzeń parametru kontrolnego po przekroczeniu wartości dopuszczalnej. Zaprezentowano przykład szacowania trwałości tłoczkowej pompy hydraulicznej z samolotu użytkowanego w Siłach Zbrojnych RP. Dla wybranych parametrów kontrolnych pompy hydraulicznej wyznaczono ich wartości dopuszczalne oraz czas pierwszej kontroli parametru kontrolnego po przekroczeniu wartości dopuszczalnej. Zaprezentowana metoda wiąże trwałość z fizycznymi mechanizmami zużywania się zespołów. Przedstawiona metoda może być wykorzystana w pracach mających na celu określanie zasobu pracy urządzeń technicznych. Umożliwia ona użytkowanie urządzeń technicznych według strategii stanu technicznego z kontrolowaniem parametrów.

Słowa kluczowe: *lotnictwo, trwałość, napęd hydrauliczny, pompa hydrauliczna, stan techniczny.*

1. Introduction

Estimating the life of aviation hydraulic drive is a broad forecasting issue at the engineering stage of their operational behaviours, as well as forecasting the change of their technical state throughout the operation stage. The experience from the operation of aircraft hydraulic propulsion in aircraft indicates that after using the normative durability established by the manufacturer, most hydraulic assemblies still have some work resource that can be used [21, 24]. This may indicate that at the design stage of hydraulic units, their operating conditions were incorrectly identified and inadequate redundancy was imposed when estimating their durability [21].

Therefore, there is a need for technical and scientific search for methods of estimating durability correcting adopted design assumptions while maintaining the functionality and effects of the hydraulic assembly. Based on available literature sources, one can draw up

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

a certain view regarding the general principles of determining the life of hydraulic assemblies, adopted by various research, scientific and production facilities [1, 14].

The current practice of estimating the durability of aviation hydraulic drive units is multi-faceted and multi-directional. The main direction of estimating durability is based on the principle that based on laboratory test data and bench tests, it is possible to assess the durability of the assembly in appropriate operating conditions [7, 23]. The second direction, supplementing the main one, is estimation of durability based on tests of operational reliability of the assembly [1,13]. Both directions use safe durability concepts for team design.

The first direction of durability estimation requires conducting long-term and costly tests of hydraulic assemblies until they become unfit [2, 7]. In this approach, at the design stage, wear tests of hydraulic units are carried out [4, 18]. These tests are carried out only in the workplace [5, 11]. They are aimed at checking the assumed hydraulic resistance of precise pairs of the tested assembly [2, 22]. These tests are conducted according to specially developed schedules for the entire unit, which usually provide for their accelerated mode and load conditions harsher than the ones recorded during operation [14]. They last until the hydraulic unit is damaged. In general, the test schedule, determined in the course of developing a given hydraulic unit or drive, includes the execution of a series of identical, subsequent stages, each of which consists of a number of sub-stages with varying values of load parameters for a given unit, conducted at a specified time, therefore, over a specified number of load cycles [9, 12]. Hence it can be seen that the study period of time is long, therefore, the tests are also expensive. However, wear tests do not take into account all operational forces, since the very reconstruction at a test bench of the loads actually occurring in the course of operation of a studied unit is a huge problem. The principles for the determination of each test schedule are also an issue, which is extremely comprehensive and time-consuming. The dispersion of the results of experimental wear tests when estimating the durability of a hydraulic unit is the basis for introducing a safety factor, i.e. an unnamed ratio of dangerous value to limit value. Most usually, a safety factor takes into account the potential inadequacy of the test schedule relative to actual operating conditions in the course of operation [2], the availability of wear location for verification [19], the nature of developing destruction and its rate [14], the degree of credibility of determining loads for a given unit [24] and the population of a test bench studied sample [2]. Agamirov and Reicher assumed the value of the coefficient taking into account the possible inadequacy of the test program to the actual working conditions of the team equal 1.0 [2], and Ignatowicz with the team equal 1.5 [14]. Taking into account the availability of the place of wear for control, the nature of the progressing destruction and the speed of destruction, Otshu and the team assumed a value of 1.2 [19]. Taking into account the sample size, in the case of one sample tested at the stand, the value of the factor of 5 was adopted, and for six samples the value of 3 [2]. Generally, the value of the safety factor is taken from 1 to 5 [14, 24]. Based on the results of wear tests and taking into account safety factors, normative durability is determined [1, 26].

Another approach for estimating the durability of hydraulic units using the concept of safe durability are operational reliability tests [17, 27]. This strategy involves using the assembly on the aircraft until damage occurs. This strategy uses statistical methods as well as computer simulation techniques and programmed reliability tests. This strategy can be used only if the consequences of damage do not violate the principles of occupational safety and do not increase the operating costs of hydraulic units [27].

Methods for estimating durability based on safety factors do not provide an opportunity to assess the function of the distribution of the durability of a hydraulic assembly at the design stage. Therefore, works are also carried out to ensure the efficient operation of hydraulic units, using modern diagnostic methods [16, 20]. The main direction of this work is the development of methodologies for prognostic management and management of the technical condition of teams based on combining many sources of information from exploitation. For the processing of operational data, modern techniques of tracking neural networks are used, as well as automatic inference algorithms and failure probability progression algorithms [6, 8]. The so-called research also falls under this trend residual stability, which uses extended Kalman filter technique models, time series prediction, multidimensional data distribution, and phase space reconstruction [9, 20]. The influence of contamination on the durability of various hydraulic pairs of precision hydraulic units is also investigated [25]. An experimental method of measuring sensitivity to pollution based on the pollution sensitivity model was used to predict the lifetime of the hydraulic assembly. In some works, the reliable operation time of a renewable technical object was identified by applying three criteria, which used the following statistics: modified Kolmogorov-Smirnov (MK-S) statistics, statistics of the average absolute deviation of the hypothetical and empirical cumulative distribution and statistics calculated on the basis of logarithmized reliability function [3, 21]. The values of these statistics were used to rank eleven damage probability distributions. It has been shown that based on the aggregate criterion taking into account three statistics of match compliance, the reliability of the estimation of the work time to damage distribution increases, thus avoiding mistakes that can be made by becoming addicted to only one of them.

Agamirov and Vestyak and Blancke with the team showed that hydraulic drives have a strong correlation between parameters determining their condition and time of use [1, 7]. Therefore, one can predict the moment of change in the technical condition of the hydraulic drive assembly, provided that this condition is periodically checked [1, 24]. Using this property, the authors of this article proposed a priori-predictive method of estimating durability.

The method presented in this article is based on the observation of a selected control parameter of the hydraulic drive assembly during its use. The purpose of this check is to detect in advance the pre-emergency (allowable) condition. The novelty of the method is the use of pre-fault detection to detect pre-fault tolerance on the selected control parameter. Pre-emptive tolerances are a set of values of the selected control parameter contained between the threshold levels and before the emergency (allowable) levels. Periodic inspection of the technical condition of the hydraulic assembly using selected control parameters enables prediction of the moment of the limit state of the hydraulic drive assembly. The quantitative wear characteristics of hydraulic units change over time and their impact on the technical condition of the aviation hydraulic drive is random. The condition for implementing the method is knowing the limit level of the hydraulic drive control parameter. The limit value of the hydraulic drive unit control parameter is determined at the stage of its construction and design. It results from the design conditions of hydraulic precision pairs (plunger pairs, distribution pairs, control pairs) and functional conditions of the entire hydraulic assembly. depends largely from the materials used and the design solution of hydraulic precision pairs and is mainly confronted with the processes of destruction of these pairs as a result of their operation. The limit value for the most important parameters of a hydraulic assembly is given by the manufacturer in his technical documentation and is a reference criterion during operation.

The presented method is based on controlling the level of permissible value (pre-emergency) of the selected control parameter and determining the relationship of this parameter with the periodicity of its checks, while ensuring a given level of reliability (determined a priori reliability of the hydraulic assembly). The permissible level of the control parameter is its value at which this parameter measured at the moment t_1 does not reach the limit point t_2 with the probability $p(t) \geq p_w$, where p_w is the assumed probability level of trouble-free operation of the unit in time $\Delta\tau = t_2 - t_1$. If the value of any hydraulic control parameter η exceeds the limit value η_{dop}, but does not exceed the limit value η_{gr}, i.e. $\eta_{dop} \leq \eta \leq \eta_{gr}$, then the hydraulic drive is considered to be in a pre-emergency condition. Reaching the admissible level by the control parameter is associated with the change of the control frequency, i.e. $\Delta\tau = t_2 - t_1$. Pre-emptive tolerance $\Delta\eta = \eta_{gr} - \eta_{dop}$ is related to the frequency of the check $\Delta\tau = t_2 - t_1$ in such a way that the implementation of the process of changing the parameter determining the technical condition of the hydraulic unit, after cutting the permissible level η_{dop} at the time worked $t_1 \leq \tau \leq t_2$, does not cross until the level t_2 level η_{gr} with probability $p(t) \geq p_w$. Reaching the limit value by any control parameter enables the identification of assemblies that may soon reach the limit state. Any control parameter reaching the η_{gr} level limit, i.e. $\eta > \eta_{gr}$, means the end of the hydraulic unit's durability, i.e. the need to stop using it. It should be added here that in the case of renewable assemblies it can be subjected to a renovation procedure.

2. Description of the hydraulic unit control parameter change process

The following markings have been adopted in the following article:

- $\eta(t)$ - random function of the control parameter,
- η_{dop} -permissible value of the control parameter at random time T_{dop},
- η_{gr} -limit value of the control parameter,
- T_1 -time when the control parameter reaches the allowable level,
- T_2 -time of checking the technical condition after exceeding the permissible level (residual durability range),
- x -random time of intersection by the random function of the permissible control parameter η_{dop} or limit parameter η_{gr}.

The following assumptions were adopted for the description of the method for estimating the life of a hydraulic power unit:

1) Changes in the value of the control parameter for hydraulic power units are continuous over time and their transition from one state to another is the result of wear processes within the tribological pairs of such units.
2) The change of a hydraulic power unit control parameter η is a random process $\eta(t)$, ongoing under the influence of a wide spectrum of operating factors.
3) Data allowing for a formal description of the random process was obtained from the bench or operational tests.
4) At the design stage, the limit level ηgr of the hydraulic drive unit control parameter $\eta(t)$ was determined. The control parameter limit value does not change during the entire lifetime of the hydraulic assembly and is an impassable reference criterion.

To be able to estimate the durability of a hydraulic drive assembly, you must have a specific form of random variable distribution in the form of a probability density function.

Fig. 1 presents changes in the one-dimensional distribution density function $\phi(\eta, t)$ of the random control parameter and distribution density function $f(\eta_{dop}, t)$ of the intersection of the residual durability field border. Density function change courses divide drive life into three areas:

1) the area where the hydraulic assembly is in full working order,
2) the pre-emergency area in which there is a close relationship between the residual tolerance value of the controlled parameter and the periodicity of checks, while ensuring a given level of integrity,
3) border area, i.e. the area where the hydraulic assembly is in a state of inability to work.

Fig. 1 shows that for detecting - in good time - a pre-emergency (acceptable) state, the relationship between the periodicity of checks $\Delta\tau = t_2 - t_1$ and the preceding tolerance (residual durability) should be determined $\Delta\eta = \eta_{gr} - \eta_{dop}$ on the controlled parameter, while ensuring a given level of integrity. The moment of checking should be selected in such a way that $\eta_{dop} < \eta(T) < \eta_{gr}$.

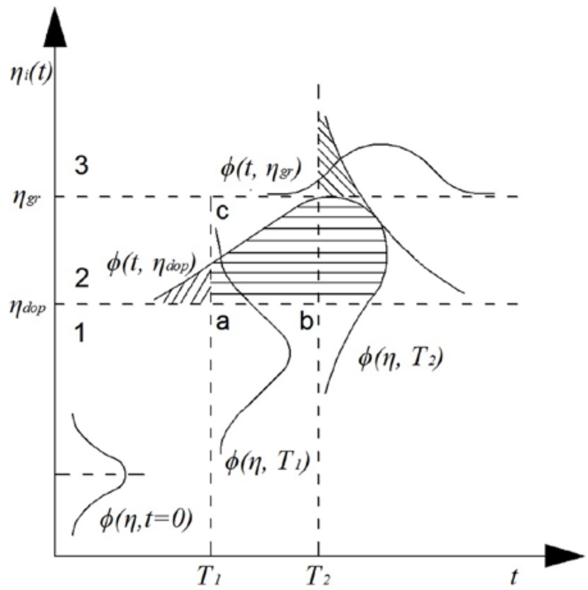


Fig. 1. The characteristics of hydraulic power unit lifetime for a random process η of this unit's control parameter change [Source: Own study]

For the control parameter level η_{dop} we have $x \leq T_1$ if and only if $\eta > \eta_{dop}$ and for the level η_{gr} we have $x \leq T_2$ if and only if $\eta \geq \eta_{gr}$. From here to the intersection of events at the level of η_{dop} we have $\{x \leq T_1\} \cap \{x \leq T_2\} = \{x \leq T_1\}$ if and only if for time T_2 we have $\{\eta > \eta_{dop}\} \cap \{\eta > \eta_{gr}\} = \{\eta > \eta_{gr}\}$. Therefore, we can note that:

$$P\{x \leq T_1\}_{\eta_{dop}} = P\{\eta > \eta_{gr}\}_{T_2},$$

which means that probability $P\{x \leq T_1\}$ at a permissible level η_{dop} is equal to the probability $P\{\eta > \eta_{gr}\}$ at moment T_2 of checking the technical condition after exceeding the permissible level. Hence:

$$\int_0^{T_1} f(x / \eta_{dop}) dx = \int_{\eta_{gr}}^{\infty} \phi(\eta / T_2,) d\eta \tag{1}$$

where:

$f(x / \eta_{dop})$ - conditional density function of the random distribution of time x, provided that the control parameter has a value of η_{dop};

$\phi(\eta/T_2)$ - conditional random density function $\eta(t)$, provided that the working time has reached the time T_2 checking the technical condition after exceeding the permissible level.

Just like for equation, (1) equation for the permissible level η_{dop} at moment T_2 is derived:

$$\int_0^{T_2} f\left(x/\eta_{dop}\right)dx = \int_{\eta_{dop}}^{\infty} \phi(\eta/T_2)d\eta \tag{2}$$

Comparing the equation (1) to the equation (2), we get:

$$\int_{T_1}^{T_2} f\left(t/\eta_{dop}\right)dt = \int_{\eta_{dop}}^{\eta_{gr}} \phi(\eta/T_2)d\eta \tag{3}$$

The notation (3) indicates that for a monotonic random process $\eta(t)$ with a specified time T_1 and known limit level value η_{gr}, it is possible to determine the next technical condition inspection deadline T_2 and the permissible level value η_{dop} at that time. The following equation results from writing the equation (3):

$$\int_{T_1}^{T_2} f\left(t/\eta_{gr}\right)dt = \int_0^{T_1} f\left(t/\eta_{dop}\right)dt \tag{4}$$

The above equation shows that a change in the value of the selected control parameter, after crossing the permissible level η_{dop} at the time worked $t_1 \le \tau < t_2$, will not cross to the time t_2 level $_{gr}$. All trajectories of the process of the random control parameter passing from the *ab* area (see Fig. 1) to the *bc* area cause a change in the frequency of checking the hydraulic assembly.

Changes in the values of selected control parameters of a hydraulic assembly occur continuously over time and the transition of the hydraulic assembly from one state to another occurs as a result of wear processes of precise tribological pairs of these assemblies. Due to the fact that the occurrence of damage to a hydraulic assembly element is caused by accidental changes in the intensity of the wear process, a linear course of the wear process can be assumed. This allows us to describe the wear process of precise tribological pairs of the hydraulic assembly by normal distribution.

Let us assume that for normal distribution, the expected value $m_\eta(t)$ and the mean quantile deviation $\sigma_\eta(t)$ are approximated linear relationships:

$$\begin{gathered} m_\eta(t) = m_a + m_b t \\ \sigma_\eta(t) = \sigma_a + \sigma_b t \end{gathered} \tag{5}$$

Constant factors m_a and m_b in relationship (5) are determined with formulas:

$$\begin{gathered} m_a = \frac{t_{i+1}\, m_\eta(t_i) - t_i\, m_\eta(t_{i+1})}{t_{i+1} - t_i} \\ m_b = \frac{m_\eta(t_{i+1}) - m_\eta(t_i)}{t_{i+1} - t_i} \end{gathered} \tag{5a}$$

Factors σ_a and σ_b are calculated using similar formulas. Moment functions $m_\eta(t)$ and $\sigma_\eta(t)$ are determined from histograms of the distribution $\phi(\eta, t_2)$ (see Fig. 2 to 4).

Density function for the distribution $\phi(\eta, t_2)$ of the random value $\eta(t)$ at moment t_2 of the technical condition inspection, after exceeding the permissible level has the form:

$$\phi(\eta/t_2) = \frac{1}{\sqrt{2\pi}\left(\sigma_a + \sigma_b t_2\right)} exp\left[-\frac{\left(\eta - m_a - m_b t_2\right)^2}{2\left(\sigma_a + \sigma_b t_2\right)^2}\right] \tag{6}$$

Based on the relationship (4), the density function for the distribution of the first intersection of the residual life level $f\left(\eta_{dop}, t\right)$ has the form:

$$f\left(t/\eta_{dop}\right) = \frac{1}{\sqrt{2\pi}\left(\sigma_a + \sigma_b t\right)} exp\left[-\frac{\left(\eta_{dop} - m_a - m_b t\right)^2}{2\left(\sigma_a + \sigma_b t\right)}\right] \frac{d}{dt}\left(\frac{\eta_{dop} - m_a - m_b t}{\sigma_a + \sigma_b t}\right) \tag{7}$$

Substituting expressions (6) and (7) to equation (3), after differentiation and necessary transformations, we get the relationship η_{dop} and $\Delta\eta = \eta_{gr} - \eta_{dop}$ for normal distribution of the parameter:

$$\eta_{dop} = \frac{\eta_{gr}\left(\sigma_a + \sigma_b T_1\right) - \left(m_b \sigma_a - m_a \sigma_b\right)\tau}{\sigma_a + \sigma_b T_1 + \sigma_b \tau} \tag{8}$$

$$\Delta\eta = \frac{\left[\left(\eta_{gr} - m_a\right)\sigma_b + m_b \sigma_a\right]\tau}{\sigma_a + \sigma_b T_1 + \sigma_b \tau} \tag{9}$$

The moment of the control parameter reaching the permissible level T_1, that is the moment of the first verification of the control parameter, can be determined using the condition of the assumed permissible level of failure-free operation P_{b_p}, as per the following expression:

$$P\{\eta < \eta \le \infty, t_1\} = \int_{\eta_{gr}}^{\infty} \phi(\eta/t_1)d\eta \le \delta_{dop} \tag{10}$$

where: $\delta_{dop} = 1 - P_{b_p}$ is the permissible damage probability.

By substituting the distribution density function $\phi(\eta, t_2)$, i.e. relationship (6) to expression (10), it is possible to determine the time of the control parameter reaching the permissible level T_1, i.e., that is the moment of the first verification of the control parameter, in the following form:

$$T_1 = \frac{\eta_{gr} - m_a - u_{P_{bp}} \sigma_a}{m_b - u_{P_{bp}} \sigma_a} \tag{11}$$

where: $u_{p_{bp}}$ is a normal distribution quantile corresponding to probability P_{b_p}.

The time of the first inspection of the hydraulic assembly as a whole (any control parameter reaching the permissible level) will be determined from the condition:

$$t_1 = \min\left(T_{1v}, T_{1p}, T_{1\delta}\right) \tag{12}$$

where T_{1V}, T_{1p}, $T_{1\delta}$ are selected control parameters of the hydraulic unit, e.g. maximum discharge pressure, volumetric efficiency coefficient, etc.

3. Estimating the rotary lifetime of a hydraulic piston pump

Rotary piston pumps with a distribution disc and adjustable output will serve as an example showing the determination of the time needed for a control parameter $\eta(t)$ to reach the permissible level (limited life range) and the time for conducting the technical condition inspection after exceeding the permissible level (monitored life range), as well as the permissible level η_{dop} of the control parameter $\eta(t)$.

The pump test procedure involves recording, among others, its such control parameters as the maximum pumping pressure p_{tmax}, volumetric efficiency factor ϑ_{vp} and the total radial clearance in piston pairs δ_p. The aforementioned parameters shall be treated as random values, i.e. $\eta_p(t_i)$, $\eta_v(t_i)$, and $\eta_\delta(t_i)$.

For fixed values of hydraulic piston pump operating time t_i of: 0 hrs, 500 hrs and 1000 hrs, each random value $\eta_i(t_i)$ has a determined empirical distribution density function $\phi(\eta_i, t_i)$, expected value m_i and mean quantile deviation σ_i. Stochastic parameters $\phi(\eta_i, t_i)$, m_i and σ_i for the control parameters, namely, maximum pumping pressure, pump volumetric efficiency factors and the total radial clearance in piston pairs were obtained following laboratory tests and verification inspections in the course of pump operation onboard an aircraft, the results of which can be found in the internal elaborations of the Air Force Institute of Technology. By substituting the values of control parameters to relationship (5a) and then the values of these coefficients to (5), we get the function of hydraulic piston pump parameter moments for the assumed pump operating time.

Histograms for distributions $\phi(\eta, t)$ and moment functions $m_\eta(t)$, $\sigma_\eta(t)$ for the maximum pressure are shown in Fig. 2, the hydraulic pump volumetric efficiency factor in Fig. 3, and the total radial clearances in hydraulic pump piston pairs in Fig. 4.

For the volumetric efficiency factor ϑ_{vp}, the hydraulic pump piston pair parameter moment functions will be:

$$m_{\eta_v}(t) = 0{,}942 - 0{,}000065 \cdot t$$

$$\sigma_{\eta_v}(t) = 0{,}024 + 0{,}000015 \cdot t$$

For the maximum pressure p_{tmax} in [Pa], the hydraulic pump piston pair parameter moment functions will be:

$$m_{\eta_p}(t) = (215{,}6 - 0{,}0031 \cdot t)10^5$$

$$\sigma_{\eta_p}(t) = (3{,}43 + 0{,}00054 \cdot t)10^5$$

For the total radial clearance in piston pairs δ_{pt} in [μm], the hydraulic pump piston pair parameter moment functions will be:

$$m_{\eta_\delta}(t) = 49{,}34 - 0{,}00973 \cdot t$$

$$\sigma_{\eta_\delta}(t) = 18{,}8 + 0{,}0012 \cdot t$$

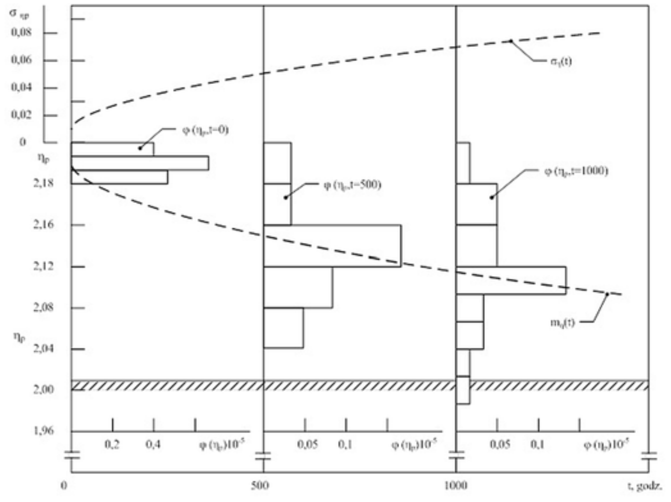


Fig. 2. Histograms for distributions $\phi(\eta, t)$ and moment functions $m_\eta(t)$, $\sigma_\eta(t)$ for the maximum pressure [Source: Own study]

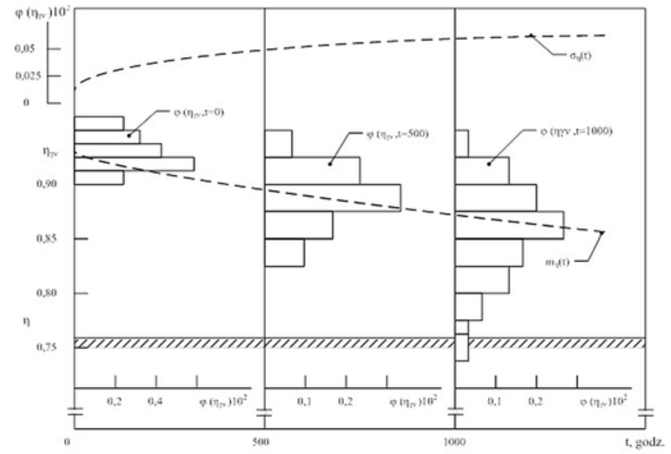


Fig. 3. Histograms for distributions $\phi(\eta, t)$ and moment functions $m_\eta(t)$, $\sigma_\eta(t)$ for hydraulic pump volumetric efficiency factor [Source: Own study]

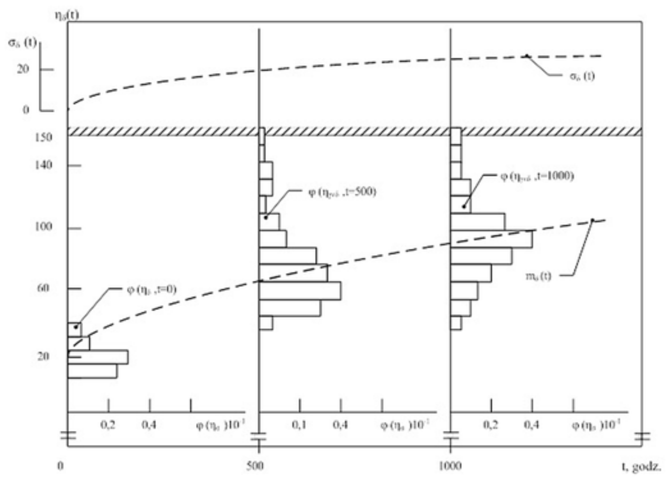


Fig. 4. Histograms for distributions $\phi(\eta, t)$ and moment functions $m_\eta(t)$, $\sigma_\eta(t)$ for hydraulic pump total radial clearances in piston pumps [Source: Own study]

Table 1. Output data for the determination of hydraulic pump control parameter moments

Parameters	η_{gri}	m_{ai}	m_{bi}	σ_{ai}	σ_{bi}
Pump volumetric efficiency factor	0,750	0,915	-0,000062	0,020	0,000012
Maximum pumping pressure [Pa]	$2000,9 \cdot 10^5$	$214,9 \cdot 10^5$	$-0,0033 \cdot 10^5$	$3,53 \cdot 10^5$	$0,00059 \cdot 10^5$
Total radial clearance in piston pairs [μm]	0,150	51,73	0,0397	18,5	0,0012

The limit levels in hydraulic pumps were determined for the volumetric efficiency factor, i.e. η_{grv}= 0.75, maximum pump pressure, i.e., η_{grp}= 200.9x10^5 Pa and the total radial clearance in piston pumps, i.e., η_{grv}= 0.150 μm. With known limit levels for control parameters and using formula (12), it is possible to determine the time for the control parameter to reach the permissible level, i.e., the moment of the first verification of the control parameter.

The output data for the determination of hydraulic pump parameter moment functions and the relationship $\eta_{dop}(t_i)$ are shown in Table 1. The verification of the hypothesis on normal distribution $\phi(\eta_i,t_r)$ using Kolmogorov's compliance test showed its compliance with optimal data.

The time for the control parameter of the pump to reach the permissible level due to its volumetric efficiency factor is $t_{1_{vp}}$ = 857 hrs, due to its maximum pressure $t_{1_{pmax}}$ = 1232 hrs, and due to its total radial clearances in piston pairs t_{1_δ} = 1326 hrs.

The time for the control parameter of the hydraulic pump to reach the permissible level shall be determined with (12):

$$t_1 = min(857,1232,1326) = 857\,\text{hrs}$$

Based on the output data shown in Tab. 1, using formula (8), it is possible to determine the relationship between the control parameter permissible level η_{dop} and the inspection periodicity for the pump parameters in question:

$$\eta_{dopv} = \frac{0,0263 + 0,00001243 \cdot \tau}{0,0268 + 0,000012 \cdot \tau},$$

$$\eta_{dopp} = \frac{(801,12 + 0,1502 \cdot \tau)10^5}{4,02 + 0,0006 \cdot \tau}\ [Pa],$$

$$\eta_{dop\delta} = \frac{2879 - 0,6879 \cdot \tau}{18,95 + 0,012 \cdot \tau}\ [\mu m].$$

Control parameter permissible value levels η_{dop} due to the pump's volumetric efficiency factor are shown in Fig. 5, due to pump's maximum pressure in Fig. 6, and due to the total radial clearance in piston pairs in Fig. 7.

The graphs presented in Fig. 5, 6 and 7 were made on the basis of calculations using the formulas (8) and (9) for functions and moments of distribution $\phi(\eta_i,t_i)$, m_i and σ_i control parameters for working time t > 500 hrs. They have they are for reference only. They present the nature of the change in the permissible level η_{dop} and the anticipating tolerance $\Delta\eta$ for the selected control parameter from the periodicity of checks τ.

For τ = 0, the allowable value of the selected control parameter reaches the limit value of this parameter, i.e. $\eta_{dop} = \eta_{gr}$ and the leading tolerance $\Delta\eta$ = 0. The end of life of the assembly is reached due to the specific control parameter. Based on the graph, e.g. the pump volume coefficient, we can determine the periodicity of checks due to

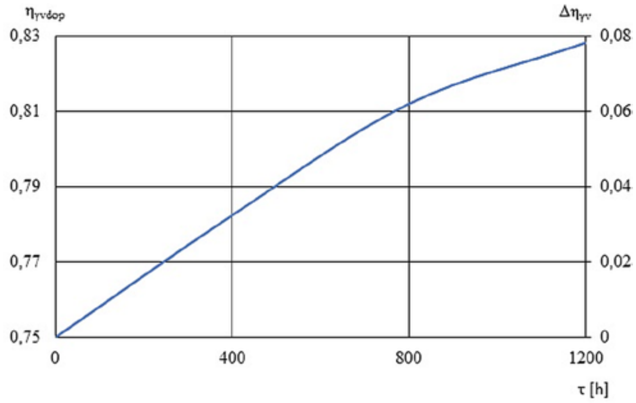


Fig. 5. Dependence of the permissible level η_{dop} and residual tolerance $\Delta\eta = \eta_{gr} - \eta_{dop}$ on verification periodicity τ the pump's volumetric efficiency factor [Source: Own study]

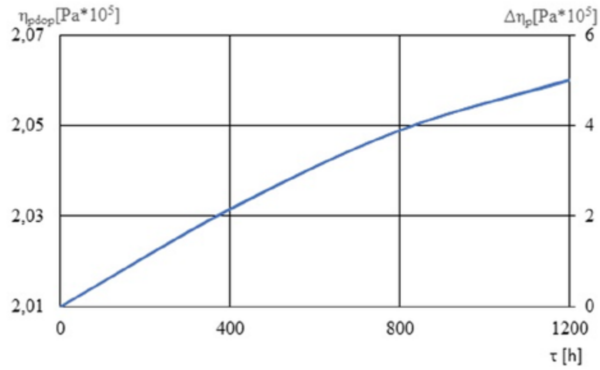


Fig. 6. Dependence of the permissible level η_{dop} and residual tolerance $\Delta\eta = \eta_{gr} - \eta_{dop}$ on verification periodicity τ for pump's maximum pressure [Source: Own study]

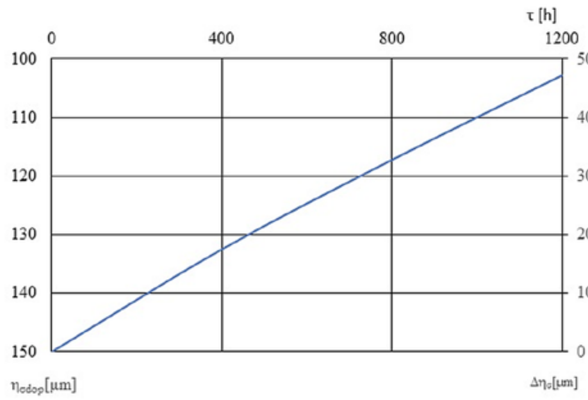


Fig. 7. Dependence of the permissible level η_{dop} and residual tolerance $\Delta\eta = \eta_{gr} - \eta_{dop}$ on verification periodicity τ for the total radial clearance in piston pairs [Source: Own study]

this parameter. If during the control the value of the volumetric efficiency coefficient will be 0.81, the time of the next inspection will be 800 hrs, while if the value of this coefficient would be 0.78, the time of the next inspection will be 400 hrs. changes the time of checking (checking).

4. Final remarks

The presented method for estimating life utilizes the property of aviation hydraulic power units, which involves a strong correlation between the parameters defining their fitness state with their operating time. It enables forecasting the hydraulic power unit limit state occurrence moment, provided that a periodic inspection of its technical condition using selected control parameters has been introduced. The purpose of this check is to detect in advance the pre-emergency (allowable) condition. In the presented method, the preceding tolerances use pre-tolerances of the selected control parameter.

The relationship between the preceding tolerance of the selected control parameter and the periodicity of its checks is presented, while ensuring the set level of a priori determined reliability of the hydraulic unit. The achievement of the pre-emergency (acceptable) level by the control parameter is associated with a change in the frequency of checks, i.e. $\Delta\tau = t_2 - t_1$ Pre-emptive tolerance size $\Delta\eta = \eta_{gr} - \eta_{dop}$ is related to the frequency of inspections $\Delta\tau = t_2 - t_1$ in such a way that the process of changing the selected control parameter that determines the technical condition of the hydraulic unit, after cutting the permissible level η_{dop} at the time worked $t_1 \le \tau < t_2$ did not intersect the level η_{gr} until t_2 with a probability not exceeding the assumed probability of trouble-free operation of the team during $\Delta\tau$. Reaching the limit value by any control parameter enables the identification of assemblies that may soon reach the limit state. If any control parameter reaches the limit level η_{gr}, i.e. $\eta \ge \eta_{gr}$, it is necessary to stop using the hydraulic assembly.

Refinement of the presented method involves binding the general relationship expressing life (fitness time) with the physical mechanisms for hydraulic unit wear and degradation of controlled elements.

To implement the method, it is necessary to specify at the design stage the limit level η_{gr} of the parameter of the controlled hydraulic drive unit $\eta(t)$.

The presented method is applied in work aimed at determining the resource life of military aircraft hydraulic drives. The method enables utilizing technical equipment according to a technical state strategy with monitoring the parameters.

References

1. Agamirov L V, Vestyak V A. Statistical analysis of results of testing aviation products in the conditions of random evaluation. Software products and systems 2017; 1 (30): 124-129, https://doi.org/10.15827/0236-235X.117.124-129.
2. Agamirov L V, Reicher V L. Fatigue life and damage to aircraft structures. Moscow: Ministry of Science and Higher Education of the Russian Federation, 2018.
3. Andrzejczak K, Selech J. Generalised Gamma Distribution in the Corrective Maintenance Prediction of Homogeneous Vehicles. Reliability and Statistics in Transportation and Communication 2019; 519-529, https://doi.org/10.1007/978-3-030-12450-2_50.
4. Ao D, Hu Z, Mahadevan S. Design of validation experiments for life prediction models. Reliability Engineering & System Safety 2017; 165: 22-33, https://doi.org/10.1016/j.ress.2017.03.030.
5. Bansal R K. Fluid Mechanics and Hydraulics Machines (Edition 9th). New Delhi: Laxmi Publications Private Limited (chapter 6), 2011.
6. Bektas O, Jones J A, Sankararaman S, Roychoudhury I, Goebel K. A neural network filtering approach for similarity-based remaining useful life estimation. The International Journal of Advanced Manufacturing Technology 2019; 101: 87-103, https://doi.org/10.1007/s00170-018-2874-0.
7. Blancke O, Tahan A, Komljenovic D, Amyot N, Claude M L. A holistic multi-failure mode prognosis approach for complex equipment. Reliability Engineering & System Safety 2018; 180: 136-151, https://doi.org/10.1016/j.ress.2018.07.006.
8. Byington C S, Watson M, Edwards D. Data driven neural network methodology to remaining life predications for aircraft actuator components. In IEEE Aerospace Conference Proceedings (IEEE Cat. No.04TH8720) 2004.
9. Chilton D. Ensuring proper maintenance and repair in projects involving and elevated hight. Fluid Power Journal 2019; 26 (9): 39-42.
10. Ge W, Wang S., Wear condition prediction of hydraulic pump. Journal of Beijing University of Aeronautics and Astronautics 2011; 37: 1410-1414.
11. Gessner T. Analizing a hydraulic system performance. Fluid Power Journal 2018; 25 (9): 12-15.
12. Grinchar N G, Sorokin P A, Karpychev V A, Sergeev K A. Analysis of change in the state of hydraulic drive of machines in operation according to the diagnostic results. Scientia Iranica B 2020; 27(1): 295 -301.
13. Idziaszek Z, Grzesik N. Object characteristics deterioration effect on task reliability - outline method of estimation and prognosis. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16 (3): 433-440.
14. Ignatowicz S R., Karuskevich M V., Maslak T P., Jutkowicze S S. Resurs and durability of aviation equipment. Kiev: Ministry of Education and Science of Ukraine, 2015.
15. Klarecki K, Hetmańczyk M P, Rabsztyn D. Influence of the selected settings of the controller on the behavior of the hydraulic servo drive. Mechatronics - Ideas for Industrial Application. Advances in Intelligent Systems and Computing 2015: 317: 91-100, https://doi.org/10.1007/978-3-319-10990-9_9.
16. Lee J, Wu F, Zhao W, Ghaffari M, Liao L, Siegel D. Prognostics and health management design for rotary machinery systems - Reviews methodology and applications. Mechanical Systems and Signal Processing 2014; 42: 314, https://doi.org/10.1016/j.ymssp.2013.06.004.
17. Li H S, Chen D N, Yao C Y. Reliability analysis of hydraulic drive system based on evidence theory and Bayesian network. Hydraulic & Pneumatics 2017; 4: 8-14.
18. Modi P N, Seth S M. Hydraulics and Fluid Mechanics Including Hydraulics Machines (19th Edition). Standard Book House, 2013.
19. Ohtsu I, Yasuda Y, Gotom H. Wear and tribological test equipment hydraulic components. Journal of Hydraulic Research 2001; 39 (2): 203-209, https://doi.org/10.1080/00221680109499821.
20. Olivares W, Vianna L, Yoneyama T. Predictive Maintenance Optimization for Aircraft Redundant Systems Subjected to Multiple Wear Profiles. Systems Journal IEEE 2018; 12 (2): 1170-1181, https://doi.org/10.1109/JSYST.2017.2667232.

21. Selech J, Andrzejczak K. An Aggregate criterion for selecting A distributionfor times to failure of components of rail vehicles. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(1): 102-111, https://doi.org/10.17531/ein.2020.1.12.
22. Srinivasan R. Hydraulic and Pneumatic Controls. Vijay Nicole Imprints Private Limited. 2/e, 2008.
23. Tomaszek H, Żurek J, Jasztal M. Prognozowanie uszkodzen zagrazajacych bezpieczenstwu lotow statkow powietrznych. Radom: Biblioteka Problemow Eksploatacji, 2008.
24. Wang H W, Teng K N. Residual life prediction for highly reliable products with prior accelerated degradation data. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(3): 379-389, https://doi.org/10.17531/ein.2016.3.9.
25. Wang X, Lin S, Wang S P. Remaining useful life prediction model based on contaminant sensitivity for aviation hydraulic piston pump. In IEEE/CSAA International Conference on Aircraft Utility Systems (AUS). 2016, https://doi.org/10.1109/AUS.2016.7748057.
26. Werbińska-Wojciechowska S. Time resource problem in logistics systems dependability modelling. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(4): 427-433.
27. Zhai Q, Chen P, Hong L, Shen L. A random-effects Wiener degradation model based on accelerated failure time. Reliability Engineering & System Safety 2018; 180: 94-103, https://doi.org/10.1016/j.ress.2018.07.003.

Leszek UŁANOWICZ
Grzegorz JASTRZĘBSKI
Paweł SZCZEPANIAK
Air Force Institute of Technology
ul. Księcia Bolesława 6
01-494 Warsaw, Poland

E-mails: leszek.ulanowicz@itwl.pl, grzegorz.jastrzebski@itwl.pl, pawel.szczepaniak@itwl.pl

Article citation info:
Kotelko M, Macdonald M, KulatungaMP, Marszalek Z. Upper-bound estimation of load-carrying capacity of perforated cold-formed thin-walled steel lipped channel columns under compression loading. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 565–573, http://dx.doi.org/10.17531/ein.2020.3.20.

Maria KOTELKO
Martin MACDONALD
Muditha P. KULATUNGA
Zuzanna MARSZALEK

UPPER-BOUND ESTIMATION OF LOAD-CARRYING CAPACITY OF PERFORATED COLD-FORMED THIN-WALLED STEEL LIPPED CHANNEL COLUMNS UNDER COMPRESSION LOADING

OSZACOWANIE GÓRNE NOŚNOŚCI PERFOROWANYCH ZIMNO FORMOWANYCH PRĘTÓW CIENKOŚCIENNYCH PODDANYCH ŚCISKANIU

Upper-bound estimation of the load-capacity of cold-formed steel sections (TWCFS) with perforations, subjected to axial compression is presented. The estimation is performed on the basis of the Yield Line Analysis ((YLA). TWCFS lipped channel sections with two sets of perforations (on the web and on the flanges) are under investigation. The comparison of experimental results, FE simulation results, European code ultimate strength predictions and upper-bound estimation based on YLA approach is carried out and presented. Some conclusions concerning an applicability of the YLA approach for ultimate strength prediction of perforated TWCFS structural members are derived.

Keywords: *cold-formed steel, thin-walled structures, load-capacity, perforations.*

W artykule przedstawiono wyniki górnego oszacowania nośności cienkościennych prętów zimno formowanych z perforacjami, poddanych osiowemu ściskaniu. Oszacowanie to jest oparte na metodzie załomów plastycznych. Rozpatrywano dwa warianty perforacji (środnika i pasów) cienkościennych prętów ceowych z żebrami końcowymi. Przeprowadzono analizę porównawczą wyników eksperymentu, wyników symulacji numerycznych MES oraz wyników obliczeń wg wzorów normatywnych normy europejskiej z wynikami oszacowania górnego nośności opartego na metodzie załomów plastycznych. Sformułowano wnioski dotyczące możliwości zastosowania tej metody do szacowania nośności cienkościennych prętów zimno formowanych z perforacjami.

Słowa kluczowe: *stal zimno formowana, konstrukcje cienkościenne, nośność, perforacje.*

1. Introduction

Nowadays, a variety of different steel products, with a large diversity of shapes, sizes, and applications are produced using the cold-forming process. The range of use of TWCFS is very wide, including buildings, the automobile industry, shipbuilding, rail transport, the aircraft industry, highway engineering, agricultural and industry equipment, office equipment, chemical, mining, petroleum, nuclear and space industries. Compared to their hot-rolled counterparts, thin-walled cold-formed steel structural members can provide more economical and efficient design solutions due to several advantages, such as a light weight, a high flexibility in obtaining various cross-sectional shapes, a highly adaptable manufacturing process with relatively little waste, and easier and faster construction. Perforations are used in many TWCFS members to accommodate services such as electrical, plumbing, and heating. They are particularly used in storage racks and in low- to mid-rise multi-storey buildings and portal frames.

Because of the thinness of section walls and the wide variety of shapes which can be used (stiffeners, lips, etc.), some specific types of buckling, which are not observed in other thin-walled structures, can appear in cold-formed steel profiles. Particularly, symmetric and anti-symmetric buckling modes, for which buckling loads are of very close values [33] as well as coupled buckling modes [4].

Thus, stability and load carrying capacity problems, related to those members, induce particular aspects which are new challenges for researchers who work on new theoretical solutions (both analytical and numerical) and it has stimulated a development in the Theory of Stability and Theory of Thin-Walled Structures, especially in recent years. The problem of the load carrying capacity of such members subjected to simple loading systems (pure bending or uniform compression) has been with satisfactory accuracy solved within the theory of thin-walled structures, as well as in design code specifications [29, 31, 39]. This theory provides sufficient methods of the load-carrying capacity evaluation, among them effective width method [12, 39], Direct Strength Method – DSM [24, 30]. Such a method should well satisfy the target reliability index in structural design [37]. Very recently, structural analyses and reliability studies that implement the system-based design-by-analysis of steel thin-walled structure (rack frames) has been presented by Cardoso et al [3]. The analysis was based on Direct Design Method (DDM), modification of DSM method.

However, there are relatively few reports of research, that have been carried out in the analysis of TWCFS members with perforations subjected to compression loading. The state of art review indicates, that the buckling and ultimate strength of such a structure under compression depends strongly on the shape, size and location of perforations [21]. Davies et al [5] proposed an analytical method of design of axially compressed perforated steel sections, based on the effective thickness approach. Kulatunga et al published results of theoretical and experimental analysis of cold-formed steel members under compression [17, 19, 20]. Moen and Schafer applied the Direct Strength Method (DSM) to the analysis of cold-formed steel columns with holes [23] and compared the results with results of experiments [25]. Inelastic buckling of perorated plates and sections under compression with perforations was investigated by Yao and Rasmussen [38]. The problem of the structural behaviour of plates with special type of perforations (cut-outs) in inelastic range was analysed, using Finite Element Method, by Falkowicz [6]. Buckling load of non-standard thin-walled channel beams with perforations, subject to bending, was investigated by Grenda [8] and Magnucka-Blandzi [23]. Also, Nedelcu [28] analysed buckling mode decomposition of thin-walled members with holes. However, the above state -of -art review shows, that reliable design specifications to predict the ultimate buckling strength of perforated steel sections are not yet available. Thus, research into an effective method to evaluate load-carrying capacity of those structures (closely related to the structure's reliability assessment) is still an open question (particularly for members undergoing local buckling). Analytical or semi-analytical methods, based on effective thickness approach, for relatively short perforated members have been not reported so far. As mentioned above, Davies [5] proposed such a method, but it was applicable for relatively long members, for which global buckling takes place.

In this paper results of numerical (finite element FE), experimental and theoretical investigations based on the Yield Line Analysis (YLA) approach associated with effective thickness approach are presented, and conclusions are drawn on this basis. The conclusions are mainly focused on the applicability of the YLA approach to the upper-bound estimate of load-carrying capacity of TWCFS members with perforations. Preliminary, selected results of the YLA analysis are presented in [22]. This approach is competitive in comparison with FE analysis or other numerical and analytical-numerical methods, since algorithms based on it are simple and calculation time is very short, in comparison with very time-consuming FE calculations. The state of art review presented above indicates, that the YLA approach has not been applied to evaluate ultimate strength of perforated thin-walled members so far.

A thin- walled structure is characterized by a certain "redundancy" of its load-carrying capacity (l-c-c). In order to evaluate this "redundancy, one has to solve the problem of stability and post-buckling behaviour of the structure, as well as the problem of failure (post-ultimate) behaviour, i.e. to estimate the upper-bound load-carrying capacity. That estimation consists in the determination of the intersection–point of a post-buckling path (evaluated using either analytical, semi-analytical method or a numerical method, e.g. Finite Element Analysis) and a rigid-plastic *post-ultimate curve* obtained from the Yield Line Analysis (YLA) [14].

Short thin-walled columns (classified in European design code as members of class 4) subjected to compression and/or bending undergo local or distortional buckling. The failure of such members is initialised by the local-global interactive buckling of plastic-elastic type, not an elastic-elastic one. Thus, the failure of those members, either in compression or bending, is always initialized by the development of a local plastic mechanism in one of the member's walls. The problem of the load carrying capacity (l-c-c) of such members subjected to simple loading systems (pure bending or uniform compression) has been satisfactorily solved within the theory of thin-walled structures, as well as in design code specifications [32]. However, the same problem for members of non-standardised sections or shapes, and, particularly with perforations, is still an open question and worthy of further investigation. One of the solutions is an appropriate upper-bound estimation of l-c-c.

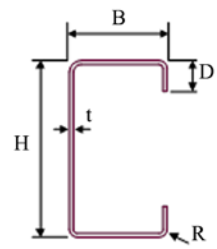
2. Subject of the analysis

The subjects of the analysis were lipped channel cross-section columns with perforations located either on the web or flanges (Figures 1 and 2). Tables 1-5 illustrate the columns dimensions, location and dimensions of perforations and material parameters taken into account in this research. The presented research is a continuation of research, the results of which (mainly experimental) have been partially published by Macdonald [22]. All investigated columns were of the same length and the same cross-section dimensions. In both sets the position of perforations (distance from the ends) was constant, but with different perforation areas (perforation diameters).

3. Numerical FE analysis

To study the structural behaviour (buckling, post-buckling and ultimate) of examined sections, the non-linear analysis was applied, using ANSYS numerical code. Detailed description of FE numerical model (geometry, boundary conditions), are given in [22]. ANSYS element type SHELL181 was used as it is adequate for linear, large rotation, and/or large strain nonlinear applications. Load was applied through a load bearing plate which represents the actual loading conditions applied in the experiments. ANSYS element type SOLID45 was used to model the load bearing plates [1] [9] [18]. Fine mesh around

Table 1. Column dimensions and material properties.

	Cross-section dimensions	
	H (mm)	120.00
	B (mm)	50.00
	D (mm)	15.00
	R (mm)	3.00
	t (mm)	1.15
Average yield strength, sy (N/mm2)		195
Modulus of elasticity, E (N/mm2)		210500

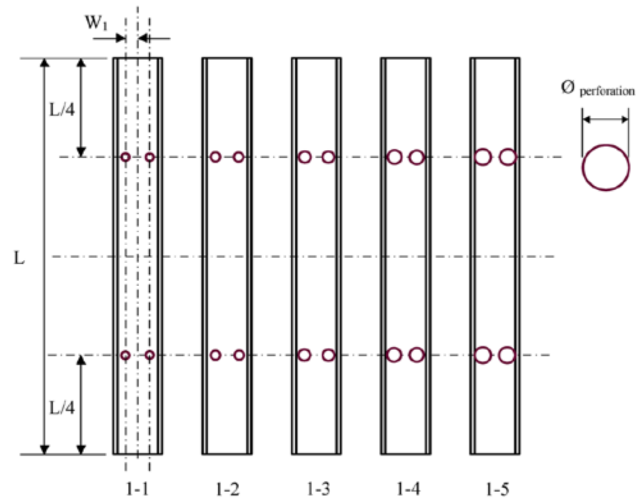


Fig. 1. Shape and location of perforations – Set 1 - perforations located on the web, [22]

Table 2. Nominal dimensions of perforations – Set 1

Column	Length of the column	Ø perforation	W1
1-1	L	3 b	c
1-2	L	4 b	c
1-3	L	5 b	c
1-4	L	6 b	c
1-5	L	7 b	c

L = 1000 mm, b = 5 mm, c = 30 mm

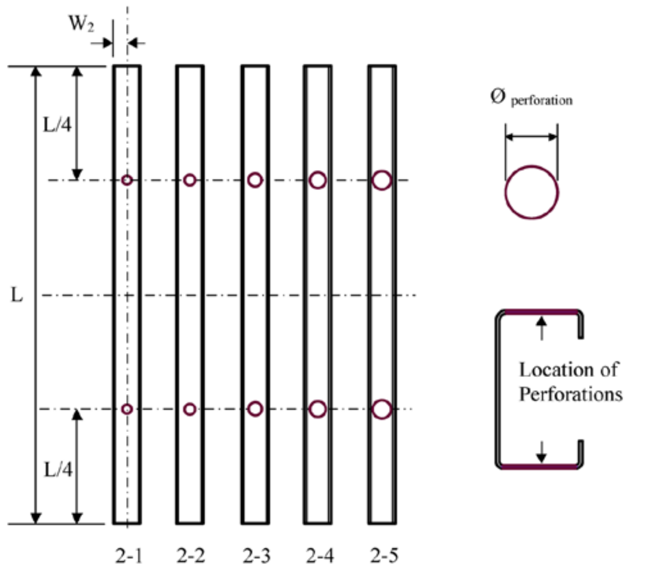


Fig. 2. Shapes and location of perforations - perforations located on the flanges – Set 2 [22]

Table 3. Nominal dimensions of perforations – Set 2

Column	Length of the column	Ø perforation	W2
2-1	L	3 b	c
2-2	L	4 b	c
2-3	L	5 b	c
2-4	L	6 b	c
2-5	L	7 b	c

L = 1000 mm, b = 5 mm, c = 25mm

Table 5. Actual measured dimensions and location of perforations

	Column Ref.	Ø perforation (mm)	Perforation position measured from bottom end (mm)
Set 1	1-1	15.02	250.02
		15.04	250.06
		15.00	750.00
		15.06	750.08
	1-2	20.80	250.04
		20.22	250.00
		20.00	750.04
		20.72	750.02
	1-3	25.02	250.00
		25.06	250.02
		25.00	750.06
		25.04	750.04
	1-4	30.02	250.02
		30.04	250.00
		30.00	750.06
		30.02	750.04
	1-5	35.04	250.02
		35.02	250.04
		35.00	250.02
		35.02	250.06
Set 2	2-1	15.02	250.02
		15.00	250.00
		15.04	750.04
		15.06	750.06
	2-2	19.98	250.02
		20.00	250.00
		20.02	750.06
		20.04	750.02
	2-3	25.00	250.02
		24.98	250.06
		25.02	250.00
		25.00	250.02
	2-4	30.02	250.02
		30.04	250.00
		30.02	250.04
		30.06	250.06
	2-5	35.02	250.02
		35.00	250.00
		35.04	750.04
		34.98	750.04

Table 4. Actual measured section dimensions

	Column Ref.	Web height, H (mm)	Flange length, B (mm)	Lip height, D (mm)	Corner radius, R (mm)	Thickness, t (mm)	Total length, L(mm)
Set 1	1-1	120.20	50.10	15.12	3.00	1.15	1000.20
	1-2	120.00	49.20	14.20	3.00	1.15	1000.00
	1-3	120.10	49.52	14.92	3.00	1.15	1000.40
	1-4	120.42	51.54	14.90	3.00	1.15	1000.10
	1-5	120.62	50.80	15.62	3.00	1.15	1000.20
Set 2	2-1	120.12	49.86	14.90	3.00	1.15	1000.30
	2-2	120.24	50.12	15.32	3.00	1.15	1000.30
	2-3	120.16	50.02	15.08	3.00	1.15	1000.30
	2-4	121.22	50.80	15.92	3.00	1.15	1000.40
	2-5	120.62	50.82	15.84	3.00	1.15	999.90

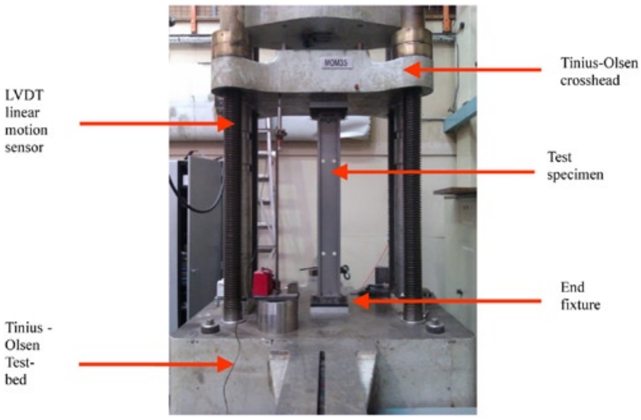


Fig. 3. Tinius olsen structural testing machine and column specimen set-up

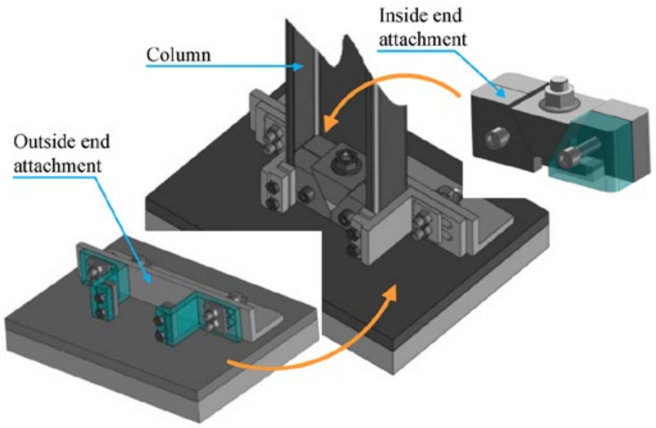


Fig. 4. Clamping grips, keeping a column specimen in a built-in position [22]

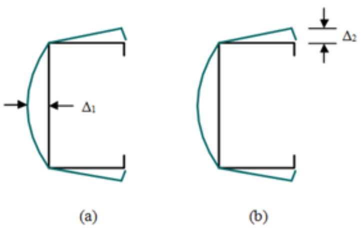


Fig. 5. Definition of geometric imperfections, (a) Type 1: geometric imperfections can be seen in a stiffened element and (b) Type 2: maximum deviation from straightness for a stiffened lip or un-stiffened flange [31]

perforations and coarse mesh further away from the perforations was applied [22]. Fixed-fixed (built-in) boundary condition where applied in FE model, adequate to actual experiments conditions.

4. Experimental Investigation

4.1. Experimental stand and specimens

All columns, with dimensions and all other testing parameters specified in Tables 1-5 were tested in the whole range of loading up to and beyond the ultimate load - to failure. All tests were carried out on the columns with fixed-fixed (built -in) boundary conditions. Experimental results reported in the next paragraphs, were used to validate Finite Element analysis results and upper-bound estimation results of the columns load-carrying capacity, based on YLA analysis. The tests were performed on the Tinius-Olsen testing machine. All column specimens were loaded with displacement control at a constant rate. A linear displacement transducer (LVDT) of high accuracy was used to measure the Tinius-Olsen testing machine crosshead displacement as shown in Figure 3.

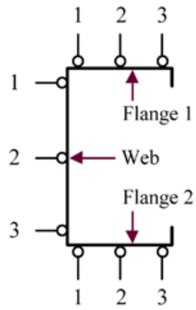


Fig. 6(a). Locations on the web and flanges to measure geometric imperfections.

Due to their comparative thinness, cold-formed thin-walled steel sections can deform easily during handling and hence, these sections can have higher initial geometric imperfections. Due to the nature of the geometric imperfections, they are categorised into two types namely, Type 1 and Type 2 as illustrated in Figure 5.

Figure 5: Definition of geometric imperfections, (a) Type 1: geometric imperfections can be seen in a stiffened element and (b) Type 2: maximum deviation from straightness for a stiffened lip or un-stiffened flange [31].

In this investigation, measurements were documented along three lines in the longitudinal direction in both web and flanges at 20 mm intervals using a co-ordinate measuring machine to a precision of 0.01 mm as presented in Figure 6. The exact values of measured dimensions are given in Tables 4 and 5.

A set of experiments was carried out to study the influence of a perforation area on the buckling and ultimate strength of column members. As mentioned above, the specimens were selected to have same cross-section dimensions, but with different perforation (Figures 2,3). The column lengths were kept constant.

4.2. Structural behaviour of tested columns

All tested columns behaved in a very similar way in whole range of loading, up to the ultimate load and beyond – in the post-ultimate stage. Buckling occurred in all specimens around 85% of the ultimate load, which corresponded to the buckling load obtained in FE calculations. All the specimens failed by local buckling, interacting with distortional buckling, at the peak load forming three local-distortional buckling half-waves. Maximum deformation occurred around the perforations near to the bottom end in the specimens set 1-2 to set 1-4 whereas, in specimens 1-1 and 1-5, the maximum deformation occurred around the perforations near to the top end. In all tested specimens the failure was initiated in the area of perforation and, subsequently, local plastic mechanisms were developed in this area (see Figure 7).

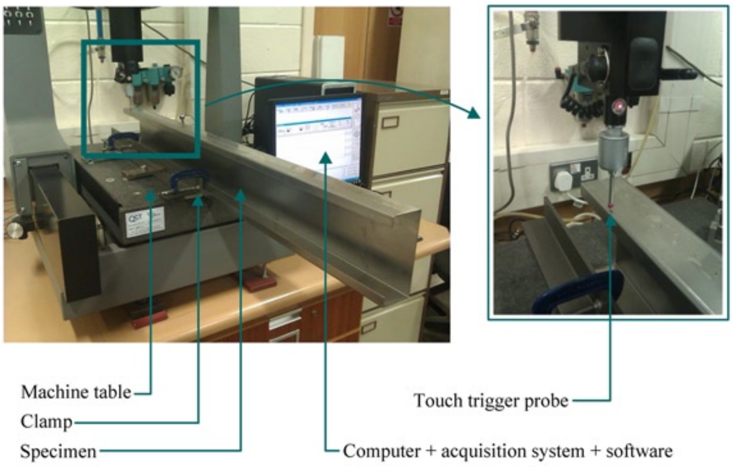


Fig. 6(b). Measuring geometric imperfections on the flange using the co-ordinate measuring machine.

5. Load-capacity upper-bound estimation via Yield Line Analysis (YLA)

5.1. YLA theoretical backgrounds

The yield line theory (Yield Line Analysis - YLA) applied to thin-walled steel structures enables an analysis of structural behaviour in the vicinity of ultimate load and in the post-ultimate stage. The basic assumption is that the plastic mechanism is fully developed, and the plastic zones developed in the walls of thin-walled steel members are concentrated at yield lines, either stationary or travelling [26, 27]. At the level of the yield lines, the material is considered to be fully plastic. Material characteristics is assumed to be rigid-perfectly plastic or rigid-plastic with strain hardening [15].

The theory distinguishes two types of the plastic mechanisms: a so-called "true mechanism" [26] and the so-called "quasi-mechanism", where the flat parts of the walls are limited by yield lines, but the walls undergo membrane deformation.

The plastic mechanism approach is based on two basic methods, namely the *energy method* (work method) and the *equilibrium strip method* [7, 16, 35]. In the present study only the energy method was applied. Using the energy method, the *Principle of Virtual Velocities* is adopted [14].

In the case of TWCFS members, the following form of this Principle, namely of the rate of change of the energy dissipated is usually applied:

$$\dot{\Pi}\left(\dot{\beta},\chi\right)=\sum_i\int_{A_i}\left(N_0\varepsilon_p\right)dA_i+\sum_j\dot{E}_{bi}\left(\dot{\beta}i,\overline{m}_{pi}\chi\right) \tag{1}$$

where Π is the potential energy of the system.

Eqn. (1) is a sum of rate of change of membrane strain energy in the walls of the global plastic hinge and bending strain energy dissipated at yield lines, where N_0 is a vector of membrane forces per unit length. The membrane forces N_0 are determined from the associated flow rule, considering corresponding yield criterion (e.g. Huber-Mises). A_i is an area of i-th plastic zone (tension field) vector, $\dot{\beta}_i$ is an angular velocity vector on the i-th yield line and $\overline{m}_{pi}$ is the plastic moment vector on that line.

The first component of Eqn. (1) is taken into account in the case of quasi-mechanisms only, which consist of both yield lines and plastic zones (tension fields). For the true mechanism only the second component of Eqn. (1) should be considered.

Rate of change of the energy dissipated at the rotation of two walls of the plastic mechanism along an i-th yield line is expressed as:

$$\dot{E}_{bi}=\int_0^{l_i}\overline{m}_{pi}\beta_i dl \tag{2}$$

For l_i = constant, the energy itself takes form:

$$E_{bi}=l_i\int_0^{\beta_i}\overline{m}_{pi}d\beta \tag{3}$$

where l_i is the length of the yield line and β_i is an angle of rotation along that line. The plastic moment in Eqn. (3) may be expressed by different equations, corresponding to different stress distributions assumed in the yield line cross-section [33]. For the fully plastic cross section, considering the rigid-perfectly plastic isotropic material model, the so called fully plastic moment takes the form [26]:

$$m_p=\frac{\sigma_y\cdot t^2}{4} \tag{4}$$

Thus, for the rigid-perfectly plastic material model, bending strain energy dissipated at i-th yield line is expressed as a product:

$$E_{bi}=l_i\cdot m_p\cdot\beta_i \tag{5}$$

After rearranging Eqn. (1) we obtain:

$$\delta W_{ext}=\delta E_b+\delta E_m \tag{6}$$

where δW_{ext} is the variation of work of external forces, δE_b is the variation of the energy of bending plastic deformation, while δE_m is the variation of the energy of membrane plastic deformation. Eqn. (6) provides a relation of generalised load (e.g. compressive load, bending moment) in terms of general displacement (e.g. shortening, the angle of rotation). The graphical representation of this relation will be termed in the present paper as a failure curve. Alternatively, Eqn. (6) may be rearranged into the following form:

$$P=\frac{\partial\left(E_b+E_m\right)}{\partial\delta} \tag{7}$$

5.2. Energy of Plastic Deformation in TWCFS Columns Subject to Compression Loading

In TWCFS columns subjected to compression the energy of plastic deformation is a sum of bending strain and membrane strain energy and generally may be expressed as follows:

$$E=E_b+\sum_{j=1}^{k}E_{mj} \tag{8}$$

where k is a number of tension fields.

The bending strain energy is given as follows:

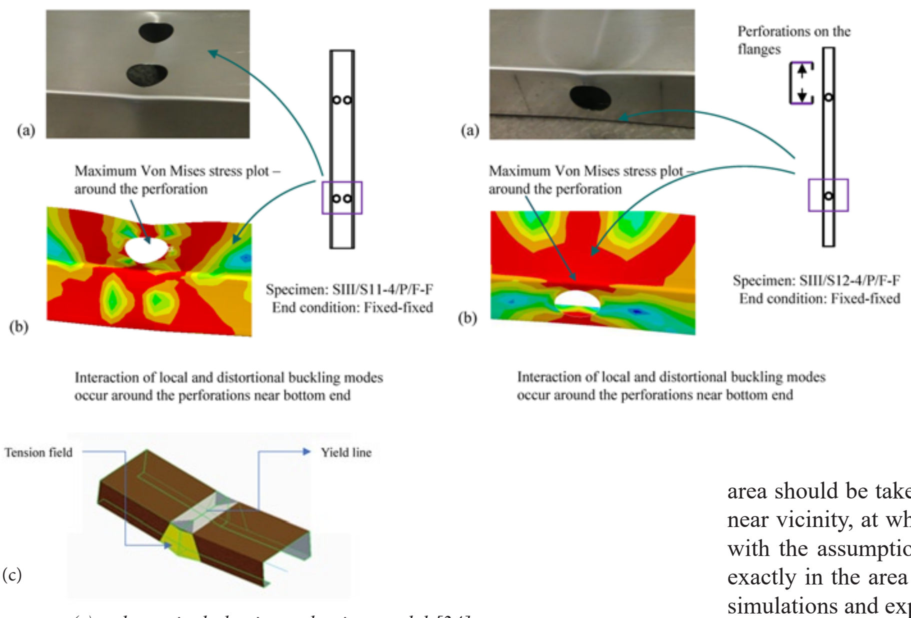
$$E_b=m_p\sum_{i=1}^{n}l_i\cdot\beta_i \tag{9}$$

The membrane strain energy E_{mj} is the energy absorbed in tension fields, either in the web or in the flange (first component in Eqn. (1)):

$$E_{mj} = \int_{A_i} \left(N_0 \varepsilon_p \right) dA_i \tag{10}$$

5.3. Theoretical Plastic Mechanism Model for Lipped-Channel Section Members with Perforations

Fig. 7. Comparison of experimental and FE deformation patterns in the post-ultimate state in the vicinity of the perforation with theoretical plastic mechanism model: (a) Set 1- perforations in the web (b) Set 2 – perforation in the flanges [22]



(c) – theoretical plastic mechanism model [34]

It was stated on the basis of experimental observations, that local plastic mechanisms were formed in all tested columns in the perforation area. The mechanism in both sets was of the same character, consisting of yield lines and tension fields, both in the web and in the flanges, so that it was classified as quasi-mechanism [26]. (see Figure 7c). The theoretical model of this mechanism was derived (shown in Figure 9), based on failure modes, observed in experiments and obtained from FE simulations (Figure7a, b). Tension fields in the model are shadowed areas (dark grey and yellow in Fig. 7c). The similar mechanism model was originally derived for TWCFS lipped channel section columns subject to eccentric compression [3]. The mechanism is a combination of pitched-roof one in the web and so called CF true mechanism in the flanges, originally derived by Murray and Khoo and described in details in [27, 35]. The mechanism shown in Fig.7, after the modification and calibration due to the numerical and experimental results, was applied in YLA calculations in the present study. The geometry of the model is shown in Figure 9 and the specification of yield lines and tension fields is given in Table 6. Since in the analysis the rigid-perfectly plastic material model was assumed, the fully plastic moment (5) was taken into account. The wall thickness t was replaced by the equivalent thickness t* in relation (4):

$$m_p = \frac{\sigma_y \cdot (t^*)^2}{4} \tag{11}$$

and in membrane energy of plastic deformation E_{mj}, given by (10), in tension fields.

The equivalent thickness was assumed to be reduced due to the perforation (t*=αt), where α is the reduction coefficient for perforated area. The reduction coefficient α was based on the ratio of the net width of a wall ($b-b_p$) to gross width b: $t^* = (1-b_p/b)t$. The weight of the net length of perforations along the member's axis (proposed by Davies [5]) was not taken into account, since in the case of YLA approach the net area should be taken for the very cross-section of perforation and its near vicinity, at which the plastic mechanism is formed. It coincides with the assumption, that the local plastic mechanism is developed exactly in the area of perforation, which was confirmed both by FE simulations and experimental tests.

Some parameters of the mechanism (among them ζ in Fig. 8 – [34]) should be determined (calibrated) on the basis of an experiment. The calibration of the mechanism model was performed for two sets of perforations separately. Within one set values of those parameters were the same. Detailed geometrical relations of this mechanism (particularly yield lines lengths l_i and rotation angles β_i in relation (4)), as well as membrane strain energy E_{mj} in tension fields (rel. 10) are given in [34].

Table 6. Specification of yield lines and tension fields in the plastic mechanism model

Yield line/ tension field	Number of yield line/ tension field	Location
AC, BD	2	Web
AE, EB, …	4	Web
EF	1	Web
ΔAEB, ΔCFD	2	Web
AG, BH	4	Flange
□AGHB, □ICJD	2	flange

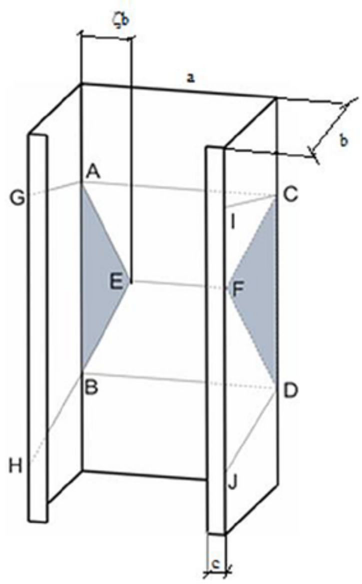


Fig. 8. Theoretical plastic mechanism model [34]; for better visibility of the figure tension fields □AGHB, □ICJD are not shadowed

5.4. Load-Carrying Capacity Upper-Bound Estimation Based on YLA

As mentioned above, upper-bound load- capacity estimation (l-c-c) is an intersection point of post-ultimate curve, obtained from (9) and post-buckling path. In the present solution, this point was obtained as an intersection of post-ultimate curve and extrapolated pre-buckling path. It is an allowable approximation, if a buckling load and ultimate load are close enough. The pre-buckling path of a column was derived, using the actual total compressive stiffness of the member with perforations. Exemplary load-

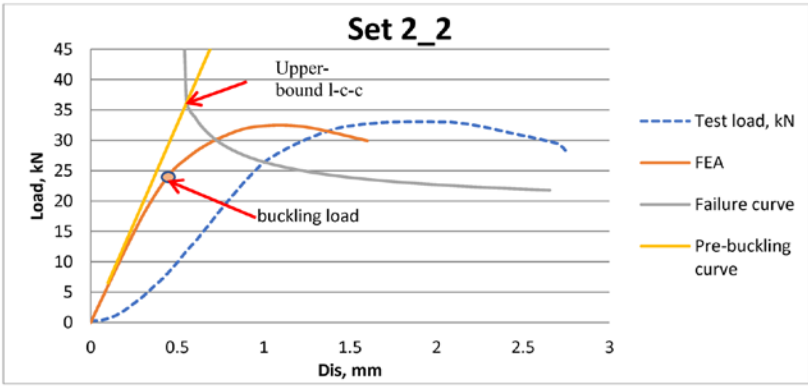


Fig. 9. Exemplary load-shortening diagra

6. Comparative analysis of numerical, theoretical and experimental result

The comparisons of numerical FE results, upper-bound l-c-c estimation results based on the analytical YLA approach and design code predictions computed from Eurocode [32] with experimental results are shown in Table 7 and 8. Since the European design code recommendations are based on ultimate load, only ultimate loads are compared. FE and test (experimental) values of ultimate loads are taken from the load-shortening diagrams, respectively.

Table 7. Ultimate strength (test results, numerical FE results, YLA results - upper-bound l-c-c, design code predictions) – Set 1

	Specimen	Ultimate Load (kN)				Ultimate Load Ratio			
		Test, $P_{exp,U}$	FEA $P_{FEA,U}$	YLA $P_{YLA,U}$	Eurocode, $N_{b,Rd}$	PFEA, U / $P_{exp,U}$	PYLA,U / $P_{exp,U}$	Nb,Rd / $P_{exp,U}$	Nb,Rd/ $P_{YLA,U}$
Set 1	1-1	36.03	35.70	35.95	40.14	0.99	1.00	1.11	1.11
	1-2	35.42	38.52	34.55	39.04	1.09	0.98	1.10	1.13
	1-3	33.56	33.99	34.52	38.55	1.01	1.03	1.15	1.12
	1-4	33.40	32.63	34.49	38.65	0.98	1.03	1.16	1.12
	1-5	33.26	32.53	34.45	38.13	0.98	1.04	1.15	1.11
					Mean, $\bar{X}$	1.01	1.02	1.13	1.12
					Standard Deviation, S	0.05	0.03	0.95	0.01
					Coefficient of Variation, COV	0.05	0.02	0.02	0.01

Table 8. Ultimate strength (test results, numerical FE results, YLA results - upper-bound l-c-c, design code predictions) – Set 2

	Specimen	Ultimate Load (kN)				Ultimate Load Ratio			
		Test, $P_{exp,U}$	FEA $P_{FEA,U}$	YLA $P_{YLA,U}$	Eurocode, $N_{b,Rd}$	$P_{FEA,U}$ / $P_{exp,U}$	$P_{YLA,U}$ / $P_{exp,U}$	$N_{b,Rd}$ / $P_{exp,U}$	Nb,Rd/ $P_{YLA,U}$
Set 2	2-1	33.14	33.50	37.4	33.41	1.01	1.13	1.01	0.89
	2-2	33.05	32.30	36.56	31.46	0.98	1.11	0.95	0.85
	2-3	30.87	30.10	35.23	28.99	0.98	1.14	0.94	0.83
	2-4	29.78	28.60	34.49	27.48	0.96	1.16	0.92	0.80
	2-5	27.04	26.31	34.45	25.13	0.97	1.27	0.93	0.73
					Mean, $\bar{X}$	0.98	1.16	0.95	0.82
					Standard Deviation, S	0.02	0.06	0.04	0.06
					Coefficient of Variation, COV	0.02	0.05	0.04	0.07

shortening diagrams (analytical pre-buckling path and post-ultimate curve) together with FEA and experimental load-shortening diagrams are shown in Figure 10. The upper bound l-c-c estimation is indicated with the arrow. In the same way upper bound l-c-c was obtained for all columns under investigation. The results (ultimate loads) are shown in Table 7 and denoted as $P_{YLA,U}$.

7. Conclusions

The comparative analysis of numerical FE simulations, experimental and analytical load-capacity prediction based on YLA method, confirms the applicability of yield line approach to the upper-bound estimation of the load-carrying- capacity of TWCFS members with perforations subject to axial compression.

This approach is competitive with FE simulations due to the simplicity of algorithms. It also may be used to perform an advanced structural analysis based on performance criteria [35].

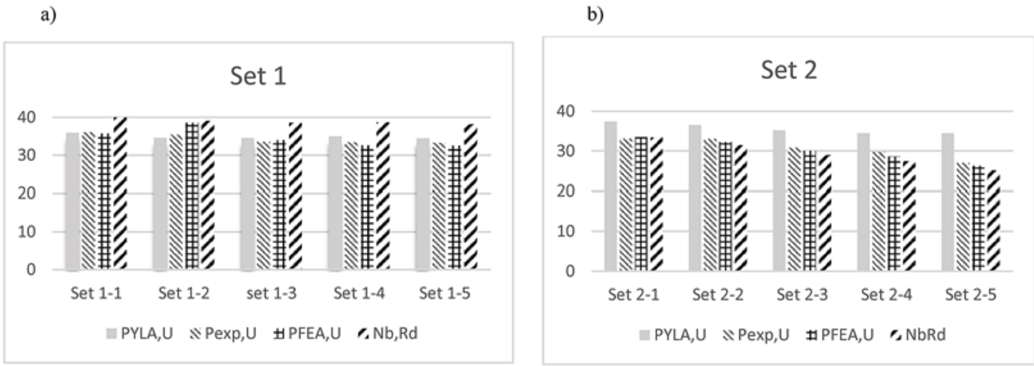


Fig. 10. Comparison of ultimate load predictions: a) set 1 – perforations in the web, b) – set 2- perforations in the flanges

The agreement of ultimate loads obtained as upper-bound estimation via YLA approach with FE and experimental results is very good for members with perforations on the web (Set 1). In that case code predictions overestimate theoretical and experimental results.

For members with perforations on the flanges the agreement of upper-bound load-carrying capacity via YLA with experimental and FE results is worse (maximum discrepancy about 25%). In that case more accurate calibration and modification of plastic mechanism model is necessary. Code predictions for that case slightly underestimate experimental and FE results and significantly underestimate load-carrying capacity via YLA approach. However, the latter is related also to the overestimation of load-capacity using the YLA method in comparison with experiment and FE simulations.

The obtained comparative results also indicated that current design rules in the European Recommendations predict not exactly the load capacity of TWCFS lipped channel column members with perforations subjected to axial compression loading and require further investigation through more extensive testing of a wider range of structural cross-sections.

References

1. ANSYS ANSYS Mechanical APDL Structural Analysis Guide: ANSYS Release 13.0. - 2010.
2. Brockenbrough R. L. Highway Engineering Handbook. McGraw-Hill Education, 2009. - ISBN: 9780071597630LCCN: 2009002381.
3. Cardoso F. System reliability-based criteria for the design of steel storage rack framesby advanced analysis: Part II - Reliability analysis and design applications. Thin-Walled Structures 2019; 141: 725-739, https://doi.org/10.1016/j.tws.2019.03.021.
4. Camotim D. Local post-buckling behavior of cold-formed steel rack columns. Proceedings. of 3th CIMS Conference (CIMS2000), Lisbon, Imperial College Press 2000: 213-222, https://doi.org/10.1142/9781848160095_0025.
5. Davies M. The design ofperforated cold-formed steel sections subject to axial load and bending. Thin-Walled Structures 1997; 29/1-4: 141-157, https://doi.org/10.1016/S0263-8231(97)00024-4.
6. Falkowicz K. Numerical analysis of compressed plates with a cut-out operating in the geometrically non-linear range. Eksploatacja i Niezawodnosc - Maintenace and Reliability 2015; 17 (2): 222-227, https://doi.org/10.17531/ein.2015.2.8.
7. Flockhart CJ, Murray NW, Grzebieta R H. Comparison of upper-bound rigid-plastic yield line mechanism analysis by the energy and equilibrium strip method. Proc. of the 2nd Australasia Congress of Applied Mechanics. Canberra, 1999.
8. Grenda M, Paczos P. Experimental and numerical study of local stability of non-standard thin-walled channel beams. Journal of Theoretical and Applied Mechanics 2019; 57: 549-562, https://doi.org/10.15632/jtam-pl/109601.
9. Hutton DV. Fundamentals of Finite Element Analysis - McGraw-Hill, 2004. - ISBN: 9780072395365LCCN: 2003048735.
10. Institution British Standards BS5950 1998: British Standards for Structural Use of Steel Work in Buildings. Part 5: Code of Practice for Design of Cold Formed Thin Gauge Sections. 1998.
11. Iron American i Institute Steel North American Specification for the Design of Cold-formed Steel Structural Members. American Iron and Steel Institute, 2016.
12. Kolakowski Z. A semi-analytical method for the analysis of the interactive buckling of thin-walled elastic structures in the second order approximation., Int. J. Solids Structures 1996; 33(25): 3779-3090, https://doi.org/10.1016/0020-7683(95)00211-1.
13. Kotełko M, Lis P, Macdonald M. Load-Capacity Probabilistic Sensitivity Analysis of Thin-Walled Beams. Proceedings of the 7th International Conference on Thin Walled Structures. Busan. 2014.
14. Kotełko M. Load-capacity and mechanisms of failure of thin-walled structures (in Polish - Nośność i mechanizmy zniszczenia konstrukcji cienkościennych), Wydawnictwa Naukowo-Techniczne, 2011.
15. Kotełko M. Load-capacity estimation and collapse analysis of thin-walled beams and columns-recent advances. Thin-Walled Structures 2004; 42: 153-175, https://doi.org/10.1016/S0263-8231(03)00055-7.
16. Kotełko M, Mania R. Alternative solutions of the problem of load-capacity of thin-walled plated structures. Mechanics and Mechanical Engineering. 2008; 2: 323-336.
17. Kulatunga M P. Load capacity of cold-formed column members of lipped channel cross-section with perforations subjected to compression loading - Part I: FE simulation and test results. Thin-Walled Structures 2014; 80: 1-12, https://doi.org/10.1016/j.tws.2014.02.017.
18. Kulatunga M, Macdonald M. Investigation of cold-formed steel structural members with perforations of different arrangements subjected to compression loading. Thin-Walled Structures 2013; 67: 78-87, https://doi.org/10.1016/j.tws.2013.02.014.
19. Kulatunga M, Macdonald M. Investigation of Cold-Formed Steel Structural Members With Perforations of Different Shapes Subjected to Compression Loading. Proceedings of the 6th International Conference on Couple Instabilities in Metal Structures. Busan 2012, https://doi.org/10.1016/j.tws.2013.02.014.

20. Kulatunga M, Macdonald M. The Efficient Design of Cold-Formed Perforated Thin-Walled Steel Structural Member Subjected to Compression Loading. Proceedings of the 7th International Conference on Thin Walled Structures. Glasgow 2014.
21. Liew JYR, Thevendran V, Shanmugam NE. Thin-Walled Structures: Research and Development. Elsevier Science, 1998.
22. Macdonald M, Kulatunga M, Kotełko M. The effects of compression loading on perforated cold-formed thin-walled steel structural members of lipped-channel cross-section. AIP Conference Proceedings 2019, https://doi.org/10.1063/1.5086138.
23. Magnucka-Blandzi E. Effective shaping of cold-formed thin-walled channel beams with double-box flanges in pure bending. Thin-Walled Structures 2011; 49: 121-128, https://doi.org/10.1016/j.tws.2010.08.013.
24. Moen C, Schafer B. Direct Strength Method for Design of Cold-Formed Steel Columns with Holes. Journal of Structural Engineering-ASCE 2010; 137, https://doi.org/10.1061/(ASCE)ST.1943-541X.0000310.
25. Moen C, Schafer B. Experiments on cold-formed steel columns with holes. Thin-walled Structures - Thin Wall Structures 2008; 46: 1164-1182, https://doi.org/10.1016/j.tws.2008.01.021.
26. Murray N W. Introduction to the Theory of Thin-Walled Structures. Clarendon Press, 1986.
27. Murray N W, Khoo P. S. Some basic plastic mechanisms in the local buckling of thin-walled steel structures. International Journal of Mechanical Sciences 1981; 23: 703-713, https://doi.org/10.1016/0020-7403(81)90008-4.
28. Nedelcu M. Buckling mode identification of perforated thin-walled members by using GBT and shell FEA. Thin-Walled Structures 2014, 82: 67-81, https://doi.org/10.1016/j.tws.2014.04.005.
29. Rhodes J. Design Cold Form Steel Members. Taylor & Francis, 1991.
30. Schafer B. Review: The Direct Strength Method of cold-formed steel member design. Journal of Constructional Steel Research 2008; 64 (7-8): 766-778, https://doi.org/10.1016/j.jcsr.2008.01.022.
31. Schafer B, Peköz T. Computational modeling of cold-formed steel: Characterizing geometric imperfections and residual stresses. Journal of Constructional Steel Research 1998; 47: 193-210, https://doi.org/10.1016/S0143-974X(98)00007-8.
32. Standardization European Committee EN 1993-1-3:2006, Eurocode 3., 2009, Design of Steel Structures; Part 1.3: General Rules - Supplementary Rules for Cold Formed Thin Gauge Members and Sheeting. - Brussels : CEN, 2009.
33. Teter A, Kołakowski Z. Lower bound estimation of load-carrying capacity of thin-walled structures with intermediate stiffeners. Thin-Walled Structures 2001; 39(8): 649-669, https://doi.org/10.1016/S0263-8231(01)00028-3.
34. Ungureanu V. et all. Plastic mechanisms of thin-walled cold-formed steel members in eccentric compression. Thin-Walled Structures 2018; 63: 183-206, https://doi.org/10.1063/1.5019077.
35. Ungureanu V, Kotełko M, Dubina D. Ultimate limit capacity of thin-walled cold-formed steel members. Ro. J. Techn. Sci. Appl. Mechanics 2018; 128: 184-192, https://doi.org/10.1016/j.tws.2017.09.029.
36. Xiao-Ling Z. Yield line mechanism analysis of steel members and connections. Progress in Structural Engineering and Materials 2003; 5: 252-262, https://doi.org/10.1002/pse.161.
37. Yan Qi L. Analysis and design reliability of axially compressed members with high-strength cold-formed thin-walled steel. Thin-Walled Structures 2007; 45: 473-492, https://doi.org/10.1016/j.tws.2007.02.012.
38. Yao Z, Rasmussen K. Inelastic local buckling behaviour of perforated plates and sections under compression. Thin-Walled Structures 2012; 61: 49-70, https://doi.org/10.1016/j.tws.2012.07.002.
39. Yu W. W. Cold-Formed Steel Design. Wiley, 2000.

Maria KOTELKO
Department of Strength of Materials
Lodz University of Technology
ul. Stefanowskiego, 90-924 Lodz, Poland

Martin MACDONALD
Muditha P. KULATUNGA
Department of Mechanical Engineering,
School of Computing, Engineering & Built Environment
Glasgow Caledonian University
Cowcaddens Rd, Glasgow G4 0BA, UK

Zuzanna MARSZALEK
Department of Strength of Materials
Lodz University of Technology
ul. Stefanowskiego, 90-924 Lodz, Poland

E-mails: maria.kotelko@p.lodz.pl, M.Macdonald@gcu.ac.uk,
muditha.kulatunga@gcu.ac.uk, jaros.zuzanna@gmail.com

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. Andrzej Dzierwa, prof. PRz
Politechnika Rzeszowska
Wydział Budowy Maszyn i Lotnictwa
35-959 Rzeszów, ul. Powstańców Warszawy 8
e-mail. adzierwa@prz.edu.pl

Dr inż. Lidia Gałda
Politechnika Rzeszowska
Wydział Budowy Maszyn i Lotnictwa
35-959 Rzeszów, ul. Powstańców Warszawy 8
e-mail. lgktmiop@prz.edu.pl

Dr hab. inż. Mirosław Tupaj, prof. PRz
Politechnika Rzeszowska
Wydział Mechaniczno - Technologiczny
37-450 Stalowa Wola, ul. Kwiatkowskiego 4
e-mail. mirek@prz.edu.pl

Dr inż. Kazimiera Dudek
Uniwersytet Rzeszowski
Centrum Innowacyjnych Technologii
35-310 Rzeszow, ul. Pigonia 1
e-mail. kaziadudek@o2.pl

Badania odporności na zużycie wybranych materiałów poddanych procesowi nagniatania ślizgowego

Słowa kluczowe: *struktura geometryczna powierzchni, tarcie, zużycie*

Streszczenie: W artykule przedstawiono wyniki badań wpływu procesu nagniatania ślizgowego realizowanego z wykorzystaniem różnych ceramik na wielkość zużycia oraz siłę tarcia elementów stalowych. Dodatkowo badaniom poddano powierzchnie po procesach szlifowania, docierania oraz polerowania. Skojarzenie materiałowe stanowiły tarcze stalowe ulepszone cieplnie do twardości 40±2 HRC oraz kulki ze stali100Cr6 o twardości 62 HRC. Badania zrealizowano przy trzech prędkościach poślizgu: 0,16 m/s, 0,32 m/s oraz 0,48 m/s. Badania udowodniły możliwość poprawy wybranych właściwości tribologicznych par trących dzięki zastosowaniu procesu nagniatania ślizgowego a także pozwoliły na ustalenie szeregu zależności pomiędzy parametrami charakteryzującymi strukturę geometryczną powierzchni oraz parametrami tribologicznymi.

Wykaz skrótów i oznaczeń:

F_{30} – wartość siły tarcia po pokonaniu drogi tarcia równej 30 m,
$F_{śr}$ – wartość średnia siły tarcia,
Hm – maksymalne wartości mikrotwardości warstwy wierzchniej,
σ_{max} – maksymalne wartości naprężeń własnych w warstwie wierzchniej,
Sal – długość autokorelacji,
Sk – wysokość rdzenia,
Sku – współczynnik nachylenia powierzchni,
Spd – gęstość wierzchołków,

Spk – zredukowana wysokość wierzchołków,
Sq – średnia kwadratowa wysokość powierzchni,
Ssk – współczynnik asymetrii powierzchni,
Svk – zredukowana głębokość wgłębień,
Sz – największa wysokość powierzchni,
VD – zużycie objętościowe tarcz,

1. Wprowadzenie

Dowolny nowy obiekt techniczny wykonany zgodnie z wymaganiami zawartymi w dokumentacji konstrukcyjnej i technologicznej posiada pełny potencjał eksploatacyjny. W miarę użytkowania i wykonanej przez niego pracy potencjał ten ulega obniżeniu w wyniku zachodzących zmian fizyko – chemicznych elementów, tj. zużycia par trących, zmęczenia materiału, procesów korozyjnych itp. [13]. Zmiany te mogą powodować obniżenie niezawodności, podwyższenie uszkadzalności czy też zmniejszenie wydajności obiektów. Wskutek oddziaływania wymuszeń roboczych przebiegają procesy niszczenia określane zużyciem [36], które powodują nagłą lub stopniową utratę właściwości użytkowych elementów. Ponieważ zużywanie w większości przypadków prowadzi do obniżenia potencjału eksploatacyjnego maszyn i ich elementów powinno się mu przeciwdziałać. Przeciwdziałanie to powinno zacząć się już na etapie konstruowania poprzez dobór odpowiednich elementów zespołu tribomechanicznego umożliwiającego redukcję zużycia w trakcie eksploatacji. Oprócz konstrukcyjnego, stosuje się również szereg sposobów technologicznych zapobiegania zużycia. Zaliczyć do nich można m.in. [1, 4, 38]:

- stosowanie obróbki cieplnej i cieplno-chemicznej (np. azotowanie utwardzające, nawęglanie, cyjanowanie),
- stosowanie obróbki plastycznej (np. młotkowanie, dogniatanie),
- stosowanie pokryć i powłok (np. niklowanie chemiczne, napawanie).

Jedną z obróbek powierzchniowych mogących opóźniać procesy zużywania jest proces nagniatania [22]. Obejmuje on m.in. takie sposoby jak rolkowanie [11], krążkowanie, nagniatanie ślizgowe [17] czy też podobne techniki jak np. shot-peening [9, 16]. W przypadku nagniatania ślizgowego twardy i gładki element nagniatający dociskany jest do powierzchni obrabianej z odpowiednią siłą wywołując w strefie nagniatania tarcie ślizgowe i w następstwie tego procesu wygładzenie powierzchni oraz korzystne zmiany właściwości warstwy wierzchniej przedmiotu [31]. Przez zastosowanie nagniatania ślizgowego można w warstwie wierzchniej podwyższyć twardość [28, 33], uzyskać bardzo dobrą gładkość powierzchni [14, 25, 30], wytworzyć naprężenia ściskające [2, 32] oraz uzyskać WW bez zanieczyszczeń ściernych. Cechy te korzystnie wpływają na szereg właściwości użytkowych, w tym na zużycie tribologiczne [5, 24], wytrzymałość zmęczeniową [33, 34] i odporność na korozję [28, 37].

W pracy [6] autorzy przeprowadzili badania nagniatania ślizgowego aluminium 6061. Wykazali oni, że poprzez zastosowanie odpowiednich parametrów obróbki możliwa jest redukcja współczynnika tarcia par trących o 48% a ubytku masy o 60-80%. Hamadache i in. [10] proces nagniatania ślizgowego zaaplikowali do stali Rb40. Badania wykazały kilkukrotny wzrost odporności na zużycie w stosunku do próbek po procesie toczenia. Wyniki badań zrealizowanych przez autorów [18] ukazują korzystny wpływ nagniatania ślizgowego na redukcję współczynnika tarcia oraz zużycia polimerów w stosunku do próbek nie nagniatanych. Badaniom poddano poliuretan oraz poliformaldehyd. Redukcja współczynnika tarcia w badaniach tribologicznych w przypadku obu polimerów sięgała maksymalnie 32%, natomiast zużycie uległo zmniejszeniu maksymalnie o 38% w stosunku do próbek po procesie toczenia. Stopień redukcji zarówno współczynnika tarcia jak i zużycia zależał mocno od wartości wyjściowej chropowatości powierzchni. Revankar i in. [26] badali właściwości tribologiczne stopu tytanu Ti-6Al-4V. Badano powierzchnie po procesie toczenia oraz po procesie nagniatania ślizgowego z różnymi parametrami wejściowymi procesu. Wykazano, że w przypadku najkorzystniejszych parametrów nastąpiła redukcja zużycia o 52% a redukcja współczynnika tarcia o 64% w stosunku do próbek toczonych. Z kolei autorzy pracy [24] badali odporność na zużycie dwufazowych stopów stali HSLA po procesie nagniatania ślizgowego. Zaobserwowano, że większą redukcją zużycia charakteryzowały się próbki nagniatane z większą siłą docisku. Celem pracy [35] było określenie możliwości ograniczenia uszkodzeń stali 316L spowodowanych zużyciem udarowym bez naruszania struktury materiału. Jako obróbkę wykończeniową powierzchni stali zastosowano nagniatanie ślizgowe. Badania wykazały redukcję objętości śladów zużycia o 53-62% w stosunku do próbek nie nagniatanych. Janczewski i in. [12] badali próbki polietylenowe o niskiej gęstości i wysokiej masie cząsteczkowej (LDPE). Obróbkę wykończeniową stanowiły procesy frezowania i nagniatania ślizgowego. Wykazano, że nagniatanie ślizgowe modyfikuje powierzchnię LDPE i radykalnie obniża wielkość zużycia, o ok. 58% w stosunku do próbek frezowanych.

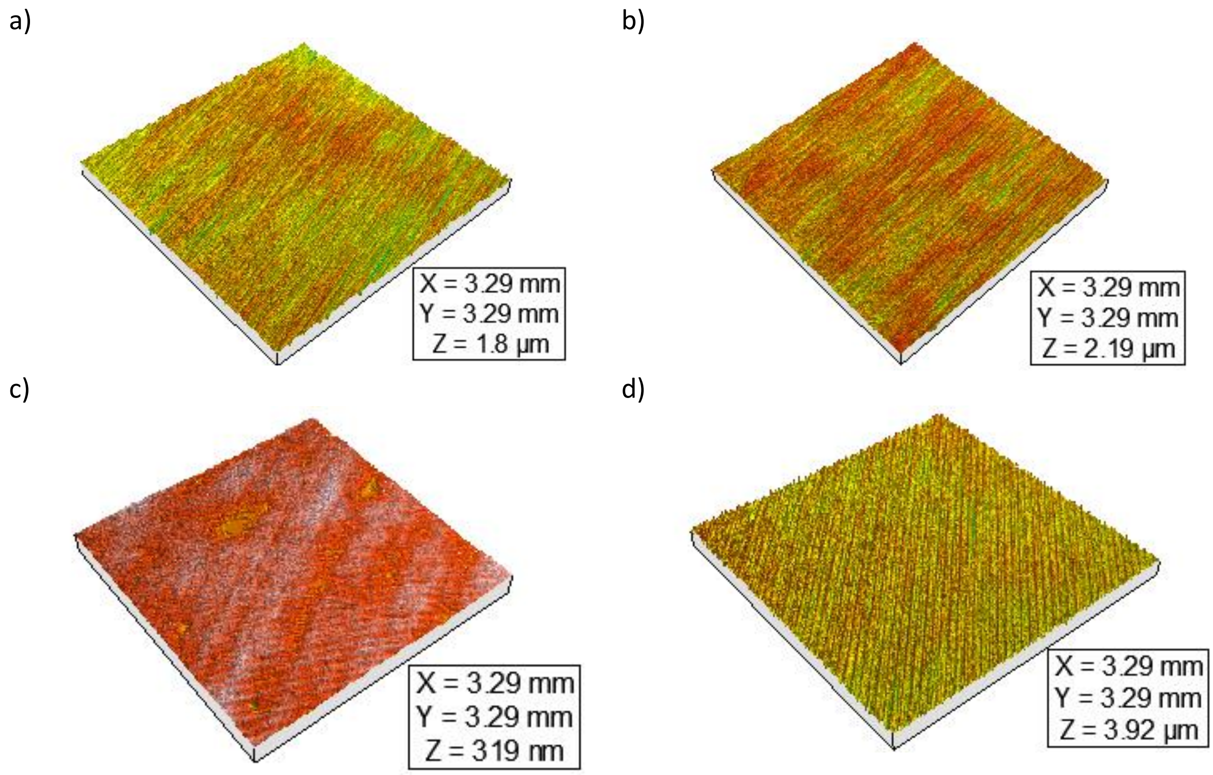
W badaniach wpływu nagniatania ślizgowego na właściwości tribologiczne skojarzeń trących, w zdecydowanej większości przypadków jako element nagniatający stosuje się diament lub kompozyt diamentowy, co sprawia, że jest to rozwiązanie stosunkowo drogie. Niewiele jest wzmianek wykorzystania tańszych elementów nagniatających aplikowanych w szczególności na obrabiarki sterowane numerycznie do poprawy tychże właściwości. Możliwość szerszego zastosowania materiałów ceramicznych [3, 26] do ich poprawy niewątpliwie prowadziłaby do redukcji kosztów związanych z procesem nagniatania w szczególności w przypadku uzyskania jakości powierzchni na poziomie podobnym do tego otrzymywanego w wyniku stosowania końcówek diamentowych. Pozwoliłaby również na wyeliminowanie lub ograniczenie pracochłonnych operacji obróbki wykończeniowej jak np. honowanie, szlifowanie, polerowanie czy też dogładzanie.

W niniejszej pracy przedstawiono badania wpływu procesu nagniatania ślizgowego realizowanego z wykorzystaniem różnych ceramik na wybrane właściwości tribologiczne skojarzeń stal – stal. Dodatkowo analizie porównawczej poddano obróbki standardowo wykorzystywane w budowie maszyn jak polerowanie, szlifowanie i docieranie.

2. Metodyka badań

Badania tribologiczne przeprowadzono na testerze tribologicznym T-11 trzpień / kulka – tarcza, w układzie kulka – tarcza. Skojarzenie tribologiczne stanowiła nieruchoma kulka łożyskowa o twardości 62±2 HRC i tarcza wykonana ze stali 42CrMo4 o twardości 40±2 HRC. Obróbkę wykończeniową tarcz stanowiły procesy nagniatania ślizgowego wykonane różnymi rodzajami kulek ceramicznych (Al_2O_3, SiC, WC). Nagniatanie wykonywano z wykorzystaniem pionowego centrum obróbczego Hass VF-3 stosując 3 różne siły docisku narzędzia: 30, 70 i 100 N. Zastosowano pojedyncze przejście narzędzia przy stałym posuwie wynoszącym 0,05 mm. Dodatkowo badaniom poddano tarcze w których obróbkę wykończeniową stanowiły procesy szlifowania, polerowania oraz docierania. Widoki izometryczne powierzchni wybranych tarcz zaprezentowano na rys. 1. Z kolei w tabeli 1 zestawiono wybrane parametry struktury geometrycznej powierzchni badanych tarcz [20]. W przypadku tarcz po procesie nagniatania ślizgowego zastosowano następujące oznaczenia:

- Al_{30}, Al_{70} oraz Al_{100} gdy elementem nagniatającym była ceramika Al_2O_3 (indeksy oznaczają zastosowaną siłę docisku przy nagniataniu),
- SiC_{30}, SiC_{70}, SiC_{100} gdy elementem nagniatającym była ceramika SiC,
- WC_{30}, WC_{70}, WC_{100} gdy elementem nagniatającym była ceramika WC.



Rys. 1. Widoki izometryczne powierzchni badanych tarcz: a) nagniatana Al_{70}, b) nagniatana SiC_{70}, c) polerowana, d) szlifowana

Tabela 1. Wybrane parametry struktury geometrycznej powierzchni badanych tarcz

	Sq	Ssk	Sku	Sz	Sal	Spd	Sk	Spk	Svk
Al_{30}	0,28	-0,898	4,36	2,84	0,0256	750	0,527	0,131	0,461
Al_{70}	0,142	-1,24	6,12	1,8	0,0172	612	0,336	0,1741	0,325
Al_{100}	0,186	-1.116	4.98	2,44	0,0185	672	0,426	0,148	0,389
SiC_{30}	0,227	-1,07	5,14	2,56	0,214	304	0,385	0,1335	0,395
SiC_{70}	0,117	-0,801	4,71	2,19	0,115	493	0,222	0,0538	0,17
SiC_{100}	0,192	-0.569	3.89	1,98	0,166	468	0,351	0,0944	0,301
WC_{30}	0,207	-0,951	5,9	3,76	0,0107	790	0,445	0,144	0,35
WC_{70}	0,3	-0,321	3,86	3,46	0,379	574	0,465	0,119	0,387
WC_{100}	0,226	-0.869	4.11	2,86	0,189	687	0,458	0,123	0,379
SZ	0,258	-0,277	3,81	3,92	0,0086	843	0,603	0,123	0,333
POL	0,0189	-0,17	2,9	0,319	0,214	1170	0,0319	0,01	0,0181
DOC	0,094	-0,59	4,39	1,3	0,0264	1020	0,198	0,0608	0,137

Dodatkowo procesy szlifowania, polerowania i docierania oznaczono odpowiednio przez „SZ”, „POL” i „DOC”. Droga tarcia we wszystkich wariantach wyniosła 282,6 m, co odpowiadało 30 minutom testu. Badania przeprowadzono przy obciążeniu 9,81 N oraz trzech prędkościach poślizgu: 0,16; 0,32 i 0,48 m/s. Podczas testów mierzono siłę tarcia, natomiast po ich przeprowadzeniu określano wielkość zużycia przy użyciu interferometru światła białego Talysurf CCI Light. Pomiarów zużycia dokonywano w czterech oddalonych od siebie o 90^0 miejscach, uzyskując obszary 3,3 mm x 3,3 mm. Następnie w każdym z nich generowano profile w kierunku prostopadłym do śladów zużycia i wykorzystując oprogramowanie TalyMap Gold 6.0 obliczano pole przekroju zużycia. Kolejnym krokiem było obliczenie zużycia objętościowego tarcz zgodnie ze wzorem (1)

$$VD = \pi d S \ [mm^3] \tag{1}$$

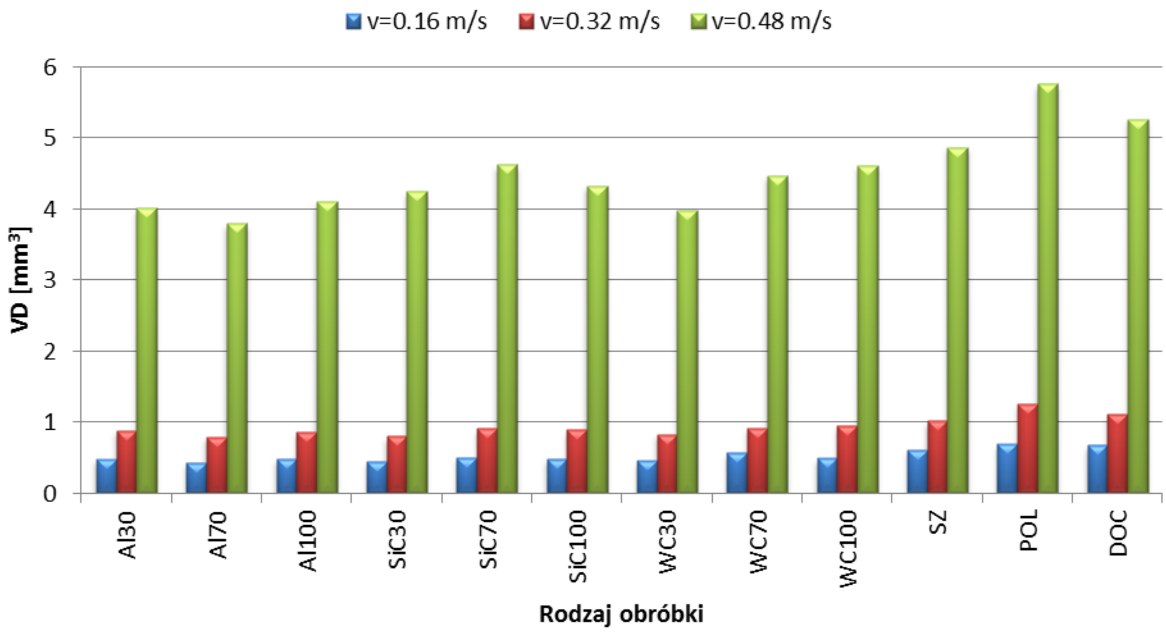
gdzie:
d – średnica tarcia [mm],
S – pole przekroju zużycia w [mm^2].

Wszystkie testy przeprowadzano z minimum trzykrotną powtarzalnością. Pomiarów naprężeń własnych dokonywano przy pomocy przenośnego dyfraktometru rentgenowskiego Xstress 3000 G3R. W pomiarach zastosowano metodę sin 2ψ [7] podczas której kąt padania ψ znajdował się w zakresie od −45° do + 45° podzielonym na 7 pozycji pochylenia. Czas ekspozycji ustawiono na 40 s. Głębokość penetracji wynosiła około 10 µm. Do pomiarów wykorzystano oprogramowanie XTronic. Dla każdej próbki wyznaczano naprężenia własne w 2 kierunkach prostopadłym i równoległym do śladów obróbkowych.

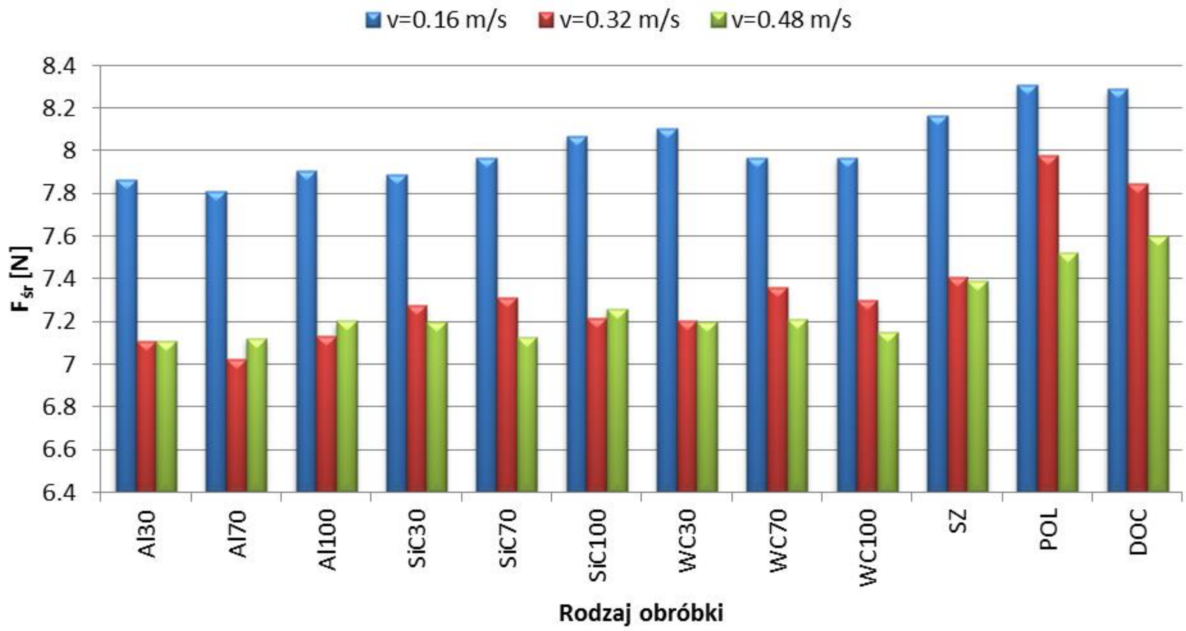
Pomiary mikrotwardości przeprowadzono na mikrotwardościomierzu Brivisor KL2 z elektroniką pomiarową HME metodą statycznego wciskania wgłębnika sposobem Vickersa [21], przy stałym obciążeniu P = 4,9 N. Czas oddziaływania wgłębnika w kształcie foremnego ostrosłupa czworokątnego o kącie dwuściennym 136° wynosił około 15 s. Mikrotwardość warstwy wierzchniej badanych próbek zmierzono na zgładach ukośnych wykonanych pod kątem 5°.

3. Prezentacja oraz omówienie wyników badań

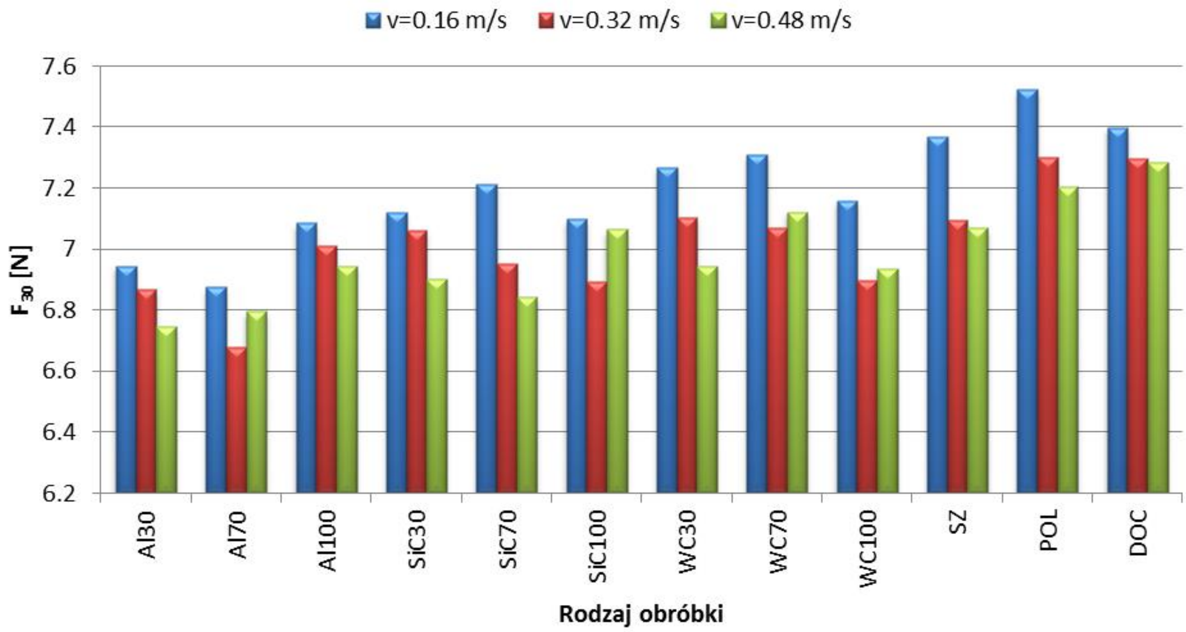
Na rysunkach 2-4 zaprezentowano wyniki otrzymanych badań. Kolejno przedstawiono średnie wartości zużycia objętościowego badanych powierzchni tarcz dla wszystkich prędkości ślizgania (rys. 2), średnie wartości siły tarcia uzyskane po procesie docierania (rys. 3) oraz wartości siły tarcia uzyskane po pokonaniu drogi tarcia równej 30 m (rys. 4). Z kolei w tabeli 2 zaprezentowano maksymalne wartości naprężeń własnych σ_{max} oraz mikrotwardości Hm a także przedziały ufności u(σmax), u(Hm) wszystkich próbek



Rys. 2. Wyniki badań zużycia objętościowego tarcz



Rys. 3. Wyniki badań średnich wartości siły tarcia



Rys. 4. Wyniki badań siły tarcia uzyskane po pokonaniu drogi tarcia równej 30 m

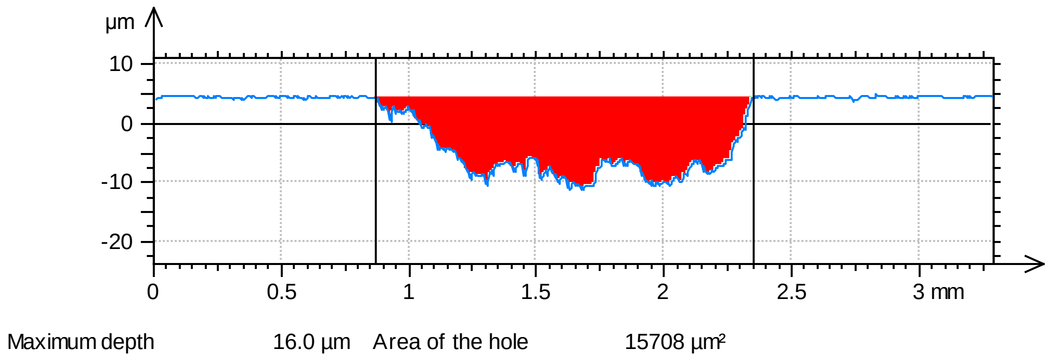
poddanych badaniom tribologicznym. Na rysunku 5 przedstawiono pola przekroju zużycia wybranych tarcz (Al_{70} oraz POL), a na rys. 6 przykładowe przebiegi siły tarcia dla tych skojarzeń przy prędkości poślizgu równej 0,16 m/s.

Przy najmniejszej prędkości poślizgu v = 0,16 m/s największe zużycie objętościowe zaobserwowano dla skojarzenia z tarczą polerowaną. Wartość parametru VD wyniosła w tym przypadku 0,694 mm^3 i korespondowała z najwyższą wartością siły tarcia uzyskaną po pokonaniu drogi tarcia równej 30 m (7,522 N) oraz najwyższą średnią wartością siły tarcia $F_{śr}$ (8,308 N). Najmniejszą wartość parametru VD przy prędkości poślizgu v = 0,16 m/s obliczono dla próbki nagniatanej Al_{70} i było to 0,439 mm^3. W przypadku tej próbki odnotowano również najmniejszą średnią wartość siły tarcia $F_{śr}$, która wyniosła 7,808 N.

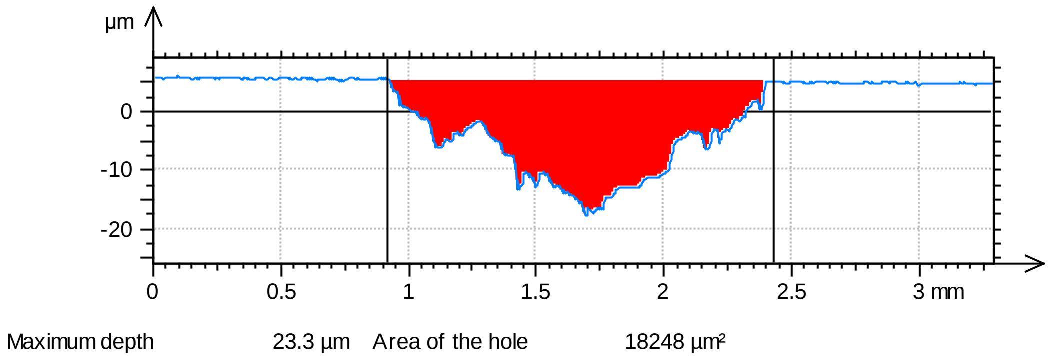
Tabela 2. Maksymalne wartości naprężeń własnych σ_{max} oraz mikrotwardości Hm powierzchni badanych tarcz

Próbka	σ_{max} [MPa]	u (σ_{max}) [MPa]	Hm [HV]	u (Hm) [HV]
Al_{30}	−212,6	±18,6	380,1	±11,6
Al_{70}	−298,1	±17,8	389,4	±14,2
Al_{100}	−309,8	±26,2	404,2	±14,8
SiC_{30}	−270,3	±18,6	399,6	±13,4
SiC_{70}	−301,6	±24,8	388,2	±11,6
SiC_{100}	−376,1	±28,6	399,7	±12,2
WC_{30}	−288,7	±26,1	382,4	±13,9
WC_{70}	−344,6	±28,7	403,5	±12,2
WC_{100}	−349,4	±24,4	398,8	±11,4
SZ	−37,2	±8,7	381,4	±12,2
POL	−179,2	±12,4	385,2	±14,6
DOC	−96,8	±10,2	384,1	±13,8

a)



b)

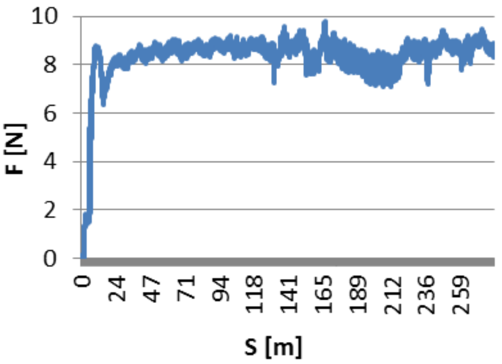


Rys. 5 Pola przekroju poprzecznego zużycia tarcz Al_{70} (a) oraz POL (b) przy prędkości poślizgu równej 0,16 m/s,

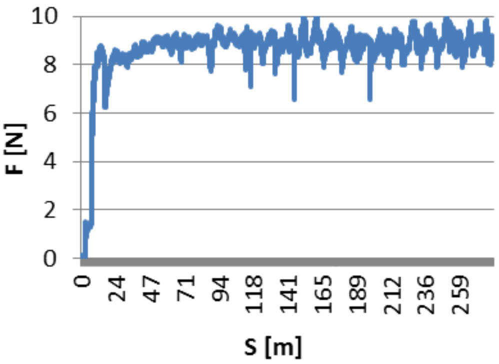
Współczynnik tarcia µ zawierał się w granicach od 0,79 (tarcza nagniatana Al_{70}) do 0,85 (tarcza polerowana).

Wzrost prędkości poślizgu do 0,32 m/s spowodował wzrost wartości zużycia objętościowego poszczególnych próbek oraz niewielki spadek średniej wartości siły tarcia oraz wartości siły tarcia uzyskanej po pokonaniu drogi tarcia równej 30 m. Podobnie jak przy najmniejszej prędkości poślizgu najmniejszą wartością zużycia objętościowego charakteryzowała się próbka Al_{70} i było to 0,796 mm^3. Również średnia wartość siły tarcia $F_{śr}$ oraz parametr F_{30} były w przypadku próbki Al_{70} najmniejsze, odpowiednio 7,022 N oraz 6,677 N. Najwyższymi wartościami poszczególnych parametrów charakteryzowała się próbka polerowana. Parametr VD w przypadku tej próbki wyniósł 1,262 mm^3. Współczynnik tarcia μ zmalał w stosunku do najmniejszej prędkości poślizgu i zawierał się w granicach od 0,72 (tarcza nagniatana Al_{70}) do 0,81 (tarcza polerowana).

a)



b)



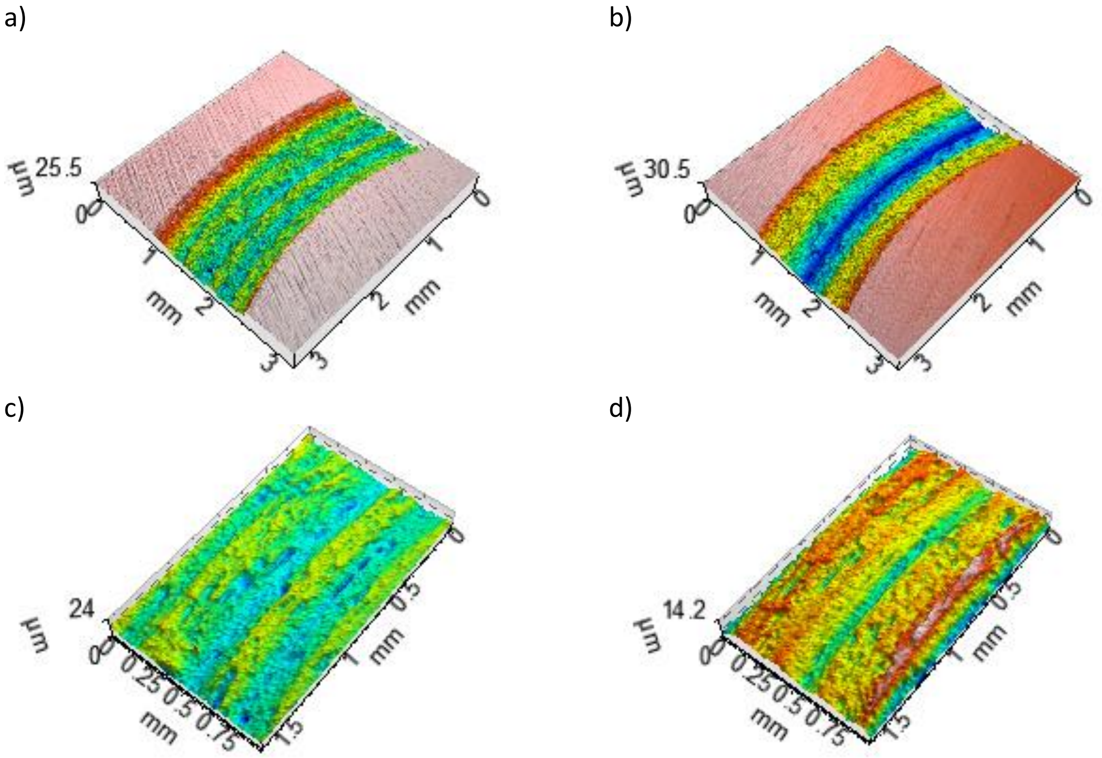
Rys. 6. Przebiegi siły tarcia par trących zawierających próbki Al_{70} (a) oraz POL (b) przy prędkości poślizgu równej 0,16 m/s

Wzrost prędkości poślizgu do 0,48 m/s spowodował bardzo wyraźny wzrost wartości zużycia objętościowego poszczególnych próbek w porównaniu do zużycia zaobserwowanego przy niższych prędkościach poślizgu. Najmniejszym zużyciem objętościowym przy tej prędkości poślizgu charakteryzowała się ponownie próbka Al_{70}, jednakże najniższe wartości $F_{śr}$ oraz F_{30} zaobserwowano w przypadku próbki Al_{30}. Parametr VD próbki Al_{70} wyniósł 3,796 mm^3. Najwyższą wartość zużycia objętościowego zmierzono ponownie w przypadku próbki polerowanej i było to 5,759 mm^3. Z kolei najwyższe wartości parametrów $F_{śr}$ oraz F_{30} odnotowano w przypadku próbki docieranej i było to odpowiednio 7,598 N oraz 7,284 N. Dolna wartość współczynnika tarcia była na podobnym poziomie jak przy prędkości 0,32 m/s natomiast górna wartość uległa niewielkiemu zmniejszeniu w stosunku do tej prędkości. Współczynnik tarcia zawierał się w granicach od 0,72 (tarcza nagniatana Al_{30}) do 0,77 (tarcza docierana).

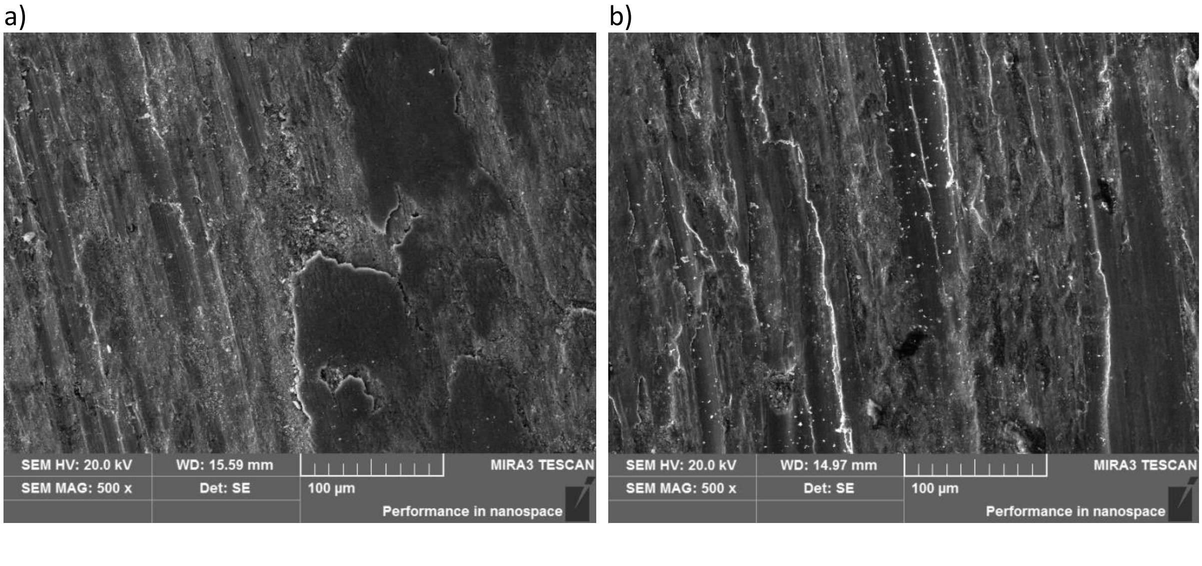
Najwyższą wartością maksymalnych naprężeń ściskających charakteryzowały się próbki nagniatane SiC_{100} oraz WC_{100}. Zmierzone wartości tych próbek to odpowiednio –376.1 MPa oraz –349.6 MPa. Z kolei najwyższą maksymalną wartością mikrotwardości powierzchni charakteryzowała się struktura geometryczna po nagniataniu kulką Al_{100}. Pomimo tego, wspomniane próbki charakteryzowały się stosunkowo dużym zużyciem objętościowym tarcz pośród próbek poddanych procesowi nagniatania ślizgowego. Najniższymi wartościami parametru VD charakteryzowały się tarcze Al_{70} i SiC_{30} przy prędkości poślizgu 0,16 m/s, ponownie tarcze Al_{70} i SiC_{30} przy prędkości poślizgu 0,32 m/s oraz tarcze Al_{70} i WC_{30} przy prędkości poślizgu 0,48 m/s.

Aby zilustrować mechanizm zużycia zużytych powierzchni, przeprowadzono analizę topografii powierzchni oraz analizę z wykorzystaniem mikroskopu skaningowego Vega3. Analiza topografii powierzchni wykazała utworzenie się jednokierunkowej tekstury po badaniach tribologicznych. Uzyskane wartości parametru Str (wskaźnik tekstury powierzchni) zawierały się w zakresie 6,98–13,06% i są to wartości charakterystyczne dla powierzchni anizotropowych po zużyciu ściernym. Rysunek 7 przedstawia widoki izometryczne tarcz Al_{70} oraz POL po badaniach tribologicznych, a także wycinki zużytych powierzchni tych tarcz. W przypadku tarczy nagniatanej Al_{70} parametr Str wyniósł 8,44%, a dla tarczy polerowanej 9,13%.

Również analiza SEM (rysunek 8) potwierdza dominujący mechanizm zużycia ściernego. W jego wyniku nastąpiła całkowita zmiana struktury geometrycznej powierzchni. Ślady zużycia na powierzchni tarcz zawierają wydłużone rozciągające się zgodnie z kierunkiem ślizgania kratery oraz wygładzone obszary z podłużnymi rowkami powstałe w wyniku deformacji plastycznej. Zaobserwowano również występowanie rozwarstwień na zużytej powierzchni dysków.

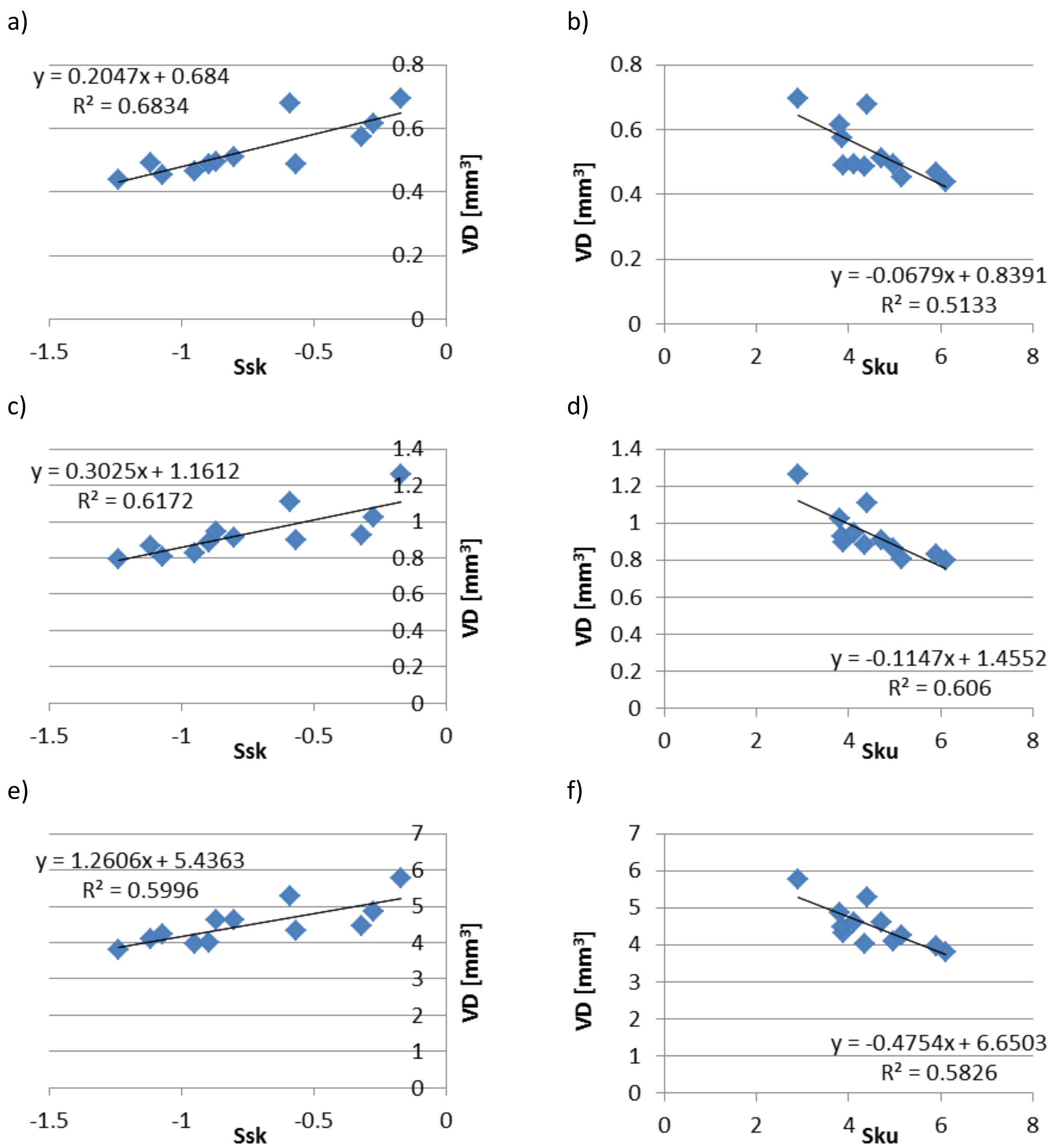


Rys. 7. Widoki izomeryczne tarcz Al_{70} (a) i POL (b) po badaniach tribologicznych oraz wycinki zużytych powierzchni tarcz Al_{70} (c) i POL (d)



Rys. 8. Zdjęcia SEM zużytych tarcz Al_{70} (a) oraz POL (b) przy prędkości ślizgania v = 0,16 m/s

Do poszukiwania związków pomiędzy parametrami tribologicznymi a parametrami charakteryzującymi strukturę geometryczną powierzchni użyto współczynnika korelacji liniowej R. Współczynnik ten zawiera się w przedziale [–1, 1]. Im większa jego wartość bezwzględna, tym silniejsza jest zależność liniowa między zmiennymi. Wartość R = 1 lub R = -1 informuje o pełnej zależności liniowej między cechami, natomiast R = 0 wskazuje na brak takiej zależności. Silną zależność liniową występującą przy wszystkich prędkościach ślizgania znaleziono dla par VD–Ssk (skośność powierzchni) oraz VD–Sku (kurtoza powierzchni). Zależności te przedstawiono na rysunku 9. Współczynnik skośności Ssk, zwany również współczynnikiem asymetrii charakteryzuje symetrię rozkładu rzędnych wysokości chropowatości względem płaszczyzny średniej [15]. Wg [27] ujemna wartość tego współczynnika wskazuje na powierzchnię o wzniesieniach płaskowyżowych, zaś dodatnia na powierzchnię o wzniesieniach w kształcie zaostrzonym. Wszystkie tarcze poddane badaniom



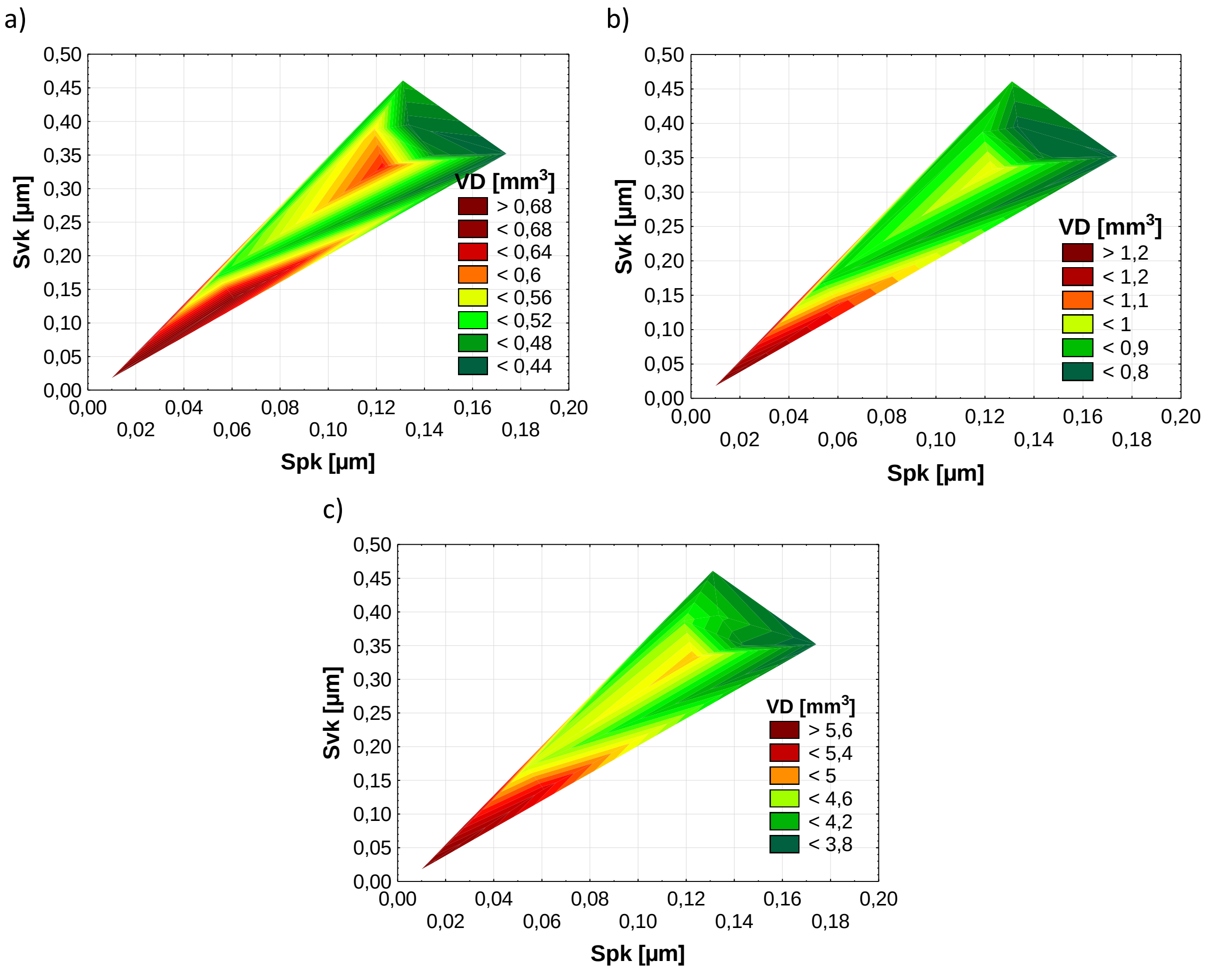
Rys. 9. Zależności pomiędzy zużyciem objętościowym VD a parametrami Ssk (a, c, e) oraz Sku (b, d, f) przy prędkościach poślizgu: a, b) v = 0,16 m/s; c, d) v = 0,32 m/s; e, f) v = 0,48 m/s

charakteryzowały się ujemną wartością parametru Ssk, który przyjmował wartości w zakresie od –0,17 (tarcza polerowana) do –1,24 (tarcza po nagniataniu Al_{70}). Z kolei parametr Sku jest określany miarą smukłości krzywej rozkładu rzędnych. Dla rozkładu normalnego rzędnych profilu Sku = 3 [8]. Próbki poddane badaniom tribologicznym charakteryzowały się rozkładem normalnym rzędnych lub rozkładem zbliżonym do normalnego. W przypadku zużycia objętościowego tarcz i parametru Ssk współczynnik korelacji liniowej R przyjmował wartości od 0,77 (dla v = 0,48 m/s) do 0,83 (dla v = 0,16 m/s), a dla VD i parametru Sku współczynnik korelacji liniowej R zawierał się między –0,72 (dla v = 0,16 m/s) i –0,78 (dla v = 0,32 m/s) i była to zależność odwrotnie proporcjonalna. Najmniejsze wartości zużycia objętościowego tarcz osiągano dla minimalnych wartości parametru Ssk oraz maksymalnych wartości parametru Sku. Ujemna skośność może poprawić warunki kontaktu przez zmniejszanie wskaźnika plastyczności, co z kolei może przyspieszyć redukcję zużycia objętościowego wraz ze wzrostem wartości parametru Sku i zmniejszeniem wartości parametru Ssk. Powierzchnie o niskiej wartości Ssk i wysokiej wartości Sku mogą stanowić „pułapki” do przechwytywania cząstek zużycia. W przypadku badań w warunkach tarcia technicznie suchego oraz w temperaturze otoczenia, cząsteczki zużycia mają zazwyczaj od ok. 10 do 100 μm. Wydaje się, że część cząstek była na tyle mała, że doliny struktury geometrycznej powierzchni zadziałały jak pułapki na część produktów zużycia i przyczyniły się do zmniejszenia intensywności zużycia. Inaczej sytuacja mogłaby wyglądać w przypadku badań w podwyższonej temperaturze a kompleksowym badaniom dotyczącym struktury produktów zużycia powstałych w realizacji procesu tarcia poświęcone są prace autorów [39, 40].

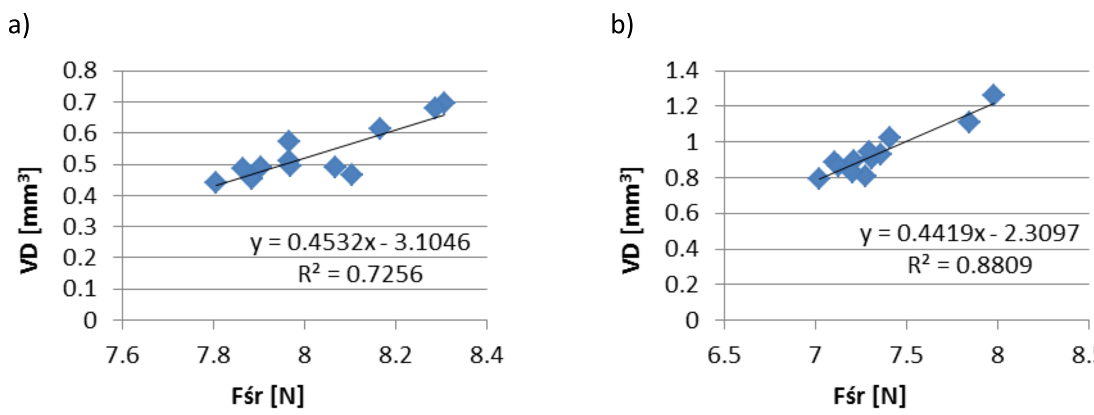
Silną zależność liniową występującą przy wszystkich prędkościach ślizgania znaleziono również dla zużycia objętościowego oraz parametrów charakteryzujących krzywą udziału materiałowego: Spk (zredukowana wysokość wzniesień), Svk (zredukowana wysokość wgłębień) i Sk (wysokość rdzenia). Zależności pomiędzy Spk i Svk zaprezentowano na rys. 10. Obszary zaznaczone kolorem czerwonym informują o większej wartości parametru VD, a kolorem zielonym (a w szczególności kolorem ciemnozielonym) o mniejszym zużyciu tarcz. Analizując otrzymane wyniki, można zauważyć, że zwiększeniu wartości parametrów Spk i Svk odpowiadało zmniejszenie zużycia objętościowego tarcz. Parametr Svk pozwala na ocenę właściwości smarnych powierzchni i jest miarą zdolności utrzymywania płynu przez powierzchnie ślizgowe. Z kolei wyższe wartości parametru Spk charakteryzują powierzchnie o wysokich szczytach, przez co obszar początkowego styku jest stosunkowo mały, a siła przyłożona na jednostkę powierzchni duża. Dlatego też parametr Spk może reprezentować nominalną wysokość materiału, który zostanie usunięty w początkowym etapie procesu eksploatacji – docierania [19]. Zwiększenie wartości Spk prowadzi jednak do zmniejszenia rzeczywistej powierzchni styku, co może ograniczać wpływ oddziaływań o charakterze adhezyjnym. Autorzy pracy [29] sugerują, że istotniejszy jest stosunek Spk/Svk, niż wartość tych parametrów oddzielnie. We wszystkich badanych strukturach geometrycznych powierzchni parametr Svk był większy niż Spk przez co skojarzenia wykazywały tendencję do redukcji współczynnika tarcia (szczególnie przy ujemnej wartości parametru Ssk). Istotna wydaje się być szczególnie zależność między parametrem Sk a zużyciem objętościowym tarczy. Parametr Sk reguluje właściwości tribologiczne obrabianych elementów po okresie docierania. W przypadku tarcz o niewielkiej wysokości nierówności ma to szczególne znaczenie gdyż okres docierania we wszystkich badanych skojarzeniach nie przekraczał 5 minut. W związku z tym parametr ten może być brany pod uwagę pod kątem planowania struktury geometrycznej powierzchni o pożądanych właściwościach tribologicznych.

Silną zależność zaobserwowano również pomiędzy zużyciem objętościowym a parametrem Spd (gęstość wierzchołków nierówności powierzchni). Parametr ten określa zagęszczenie wzniesień nierówności na badanej powierzchni. Im większa wartość parametru Spd, tym większa powierzchnia nośna [23], w szczególności w połączeniu z ujemną wartością parametru Ssk.

Mocną zależnością liniową charakteryzowało się zużycie objętościowe tarcz oraz średnia wartość siły tarcia (rys. 11). W zależności od prędkości poślizgu współczynnik



Rys. 10. Zależności pomiędzy zużyciem objętościowym VD a wartościami parametrów Spk i Svk przy prędkościach poślizgu: v = 0,16 m/s (a), v = 0,32 m/s (b), v = 0,48 m/s (c)



Rys. 11. Zależności pomiędzy zużyciem objętościowym VD a średnią wartością siły tarcia $F_{śr}$ przy prędkościach poślizgu: v = 0,16 m/s (a) oraz v = 0,32 m/s (b)

korelacji liniowej wynosił: R = 0,85 przy v = 0,16 m/s; R = 0,94 przy v = 0,32 m/s oraz R = 0,84 przy v = 0,48 m/s. Również wartości siły tarcia uzyskane po pokonaniu drogi tarcia równej 30 m były silnie skorelowane ze zużyciem objętościowym tarcz. W tym przypadku w zależności od prędkości poślizgu współczynnik korelacji liniowej pomiędzy VD a F_{30} wynosił: R = 0,85 przy v = 0,16 m/s; R = 0,73 przy v = 0,32 m/s oraz R = 0,78 przy v = 0,48 m/s. Zużycie objętościowe tarcz nie było z kolei znacząco skorelowane z maksymalną wartością naprężeń własnych (współczynnik korelacji liniowej w zakresie: -0,54 ÷ -0,64) oraz mikrotwardością (współczynnik korelacji liniowej w zakresie: -0,25 ÷ -0,34).

4. Wnioski

Na podstawie przeprowadzonych badań można stwierdzić, że wzrost prędkości poślizgu prowadził do wzrostu zużycia objętościowego tarcz poddanym badaniom tribologicznym. Spośród wszystkich próbek najniższe wartości parametru VD otrzymano w przypadku tarcz nagniatanych ślizgowo Al_{70} oraz SiC_{30}. Zużycie objętościowe tarcz było skorelowane ze średnią wartością siły tarcia a także z wartością siły tarcia otrzymaną po pokonaniu drogi tarcia równej 30 m.

Stwierdzono również występowanie szeregu zależności pomiędzy parametrami charakteryzującymi strukturę geometryczną powierzchni tarcz a zużyciem objętościowym. Silną liniową zależność zaobserwowano w przypadku parametrów charakteryzujących krzywą udziału materiałowego (Sk, Svk, Spk) a także w przypadku współczynników asymetrii oraz nachylenia powierzchni (Ssk i Sku).

Przeprowadzone badania potwierdziły korzystny wpływ procesu nagniatania ślizgowego (niezależnie od zastosowanej ceramiki) na redukcję zużycia objętościowego skojarzeń trących w porównaniu do innych popularnych obróbek wykończeniowych stosowanych w budowie maszyn.

Literatura

1. Bara M, Skoneczny W, Kaptacz S. A Tribological properties of ceramic-carbon surface layers obtained in electrolytes with a different graphite content. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2008; 4: 66-70.
2. Chomienne V, Valiorgue F, Rech J, Verdu C. Influence of ball burnishing on residual stress profile of a 15-5PH stainless steel. CIRP Journal of Manufacturing Science and Technology 2016; 13: 90-96.
3. Dzierwa A, Markopoulos A P. Influence of ball-burnishing process on surface topography parameters and tribological properties of hardened steel. Machines 2019; 7(1): 11.
4. Dzierwa A, Pawlus P, Reizer R. The effect of ceramic tribo-elements on friction and wear of smooth steel surfaces. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology 2019; 233(3): 456–465.
5. El-Tayeb N S M. Frictional behaviour of burnished copper surfaces under dry contact conditions. Engineering Research Bulletin 1994; HU Cairo: 171-184.
6. El-Tayeb N S M, Low K O, Brevern P V. Enhancement of surface quality and tribological properties using ball burnishing process. Machining Science and Technology 2008; 12: 234-248.
7. Fitzpatric M E, Fry A T, Holdway P, Kandil F A, Shackleton J, Suominen L L. Determination of Residual Stresses by X-ray Diffraction – Issue 2. A National Measurement Good Practice Guide No. 52: National Physical Laboratory, 2005.

8. Gadelmawla E S, Koura M M, Maksoud T M A, Elewa I M, Soliman H H. Roughness parameters. Journal of Materials Processing Technology 2002; 123: 133-145.
9. Ganesh B K C, Sha W, Ramanaiah N, Krishnaiah A. Effect of shot peening on sliding wear and tensile behavior of titanium implant alloys. Materials & Design 2014; 56: 480-486.
10. Hamadache H, Laouar L, Zeghib N E, Chaoui K. Characteristics of Rb40 steel superficial layer under ball and roller burnishing. Journal of Materials Processing Technology 2006; 180: 130–136.
11. Hassan A M. The effects of ball and roller burnishing on the surface roughness and hardness of some non-ferrous metals. Journal of Materials Processing Technology 1997; 72: 385–391.
12. Janczewski Ł, Toboła D, Brostow W, Czechowski K, Hagg Lobland K E, Kot M, Zagórski K. Effects of ball burnishing on surface properties of low density polyethylene. Tribology International 2016; 93: 36-42.
13. Jedliński R. Diagnostyka w procesie wyceny wartości pojazdów samochodowych i ustalaniu przyczyn oraz następstw uszkodzeń. Diagnostyka 2005, 33: 65-70.
14. Jerez-Mesa R, Gomez-Gras G, Travieso-Rodriguez J A. Surface roughness assessment after different strategy patterns of ultrasonic ball burnishing. Procedia Manufacturing 2017; 13: 710-717.
15. Jerez-Mesa R, Travieso-Rodríguez J A, Landon Y, Dessein G, Lluma-Fuentes J, Wagner V. Comprehensive analysis of surface integrity modification of ball-end milledTi-6Al-4V surfaces through vibration-assisted ball burnishing. Journal of Materials Processing Technology 2019; 267: 230-240.
16. Korzynski M, Dzierwa A, Pacana A, Cwanek J. Fatigue strength of chromium coated elements and possibility of its improvement with ball peening. Surface and Coatings Technology 2009; 204: 615–620.
17. Kovács Z F, Viharos Z J, Kodácsy J. Determination of the working gap and optimal machining parameters for magnetic assisted ball burnishing. Measurement 2018; 118: 172–180.
18. Low K O, Wong K J. Influence of ball burnishing on surface quality and tribological characteristics of polymers under dry sliding conditions. Tribology International 2011; 44: 144-153.
19. Niemczewska-Wojcik M. Wear mechanisms and surface topography of artificial hip joint components at the subsequent stages of tribological tests. Measurement 2017; 107: 89-98.
20. PN-EN ISO 25178-2:2012 Geometrical Product Specifications (GPS) – Surface texture: Areal – Part 2: Terms, definitions and surface texture parameters.
21. PN-EN ISO 4516:2004 Powłoki metalowe i inne nieorganiczne. Badania mikrotwardości metodą Vickersa i Knoopa.
22. Priyadarsini Ch, Venkata Ramana V S N, Aruna Prabha K, Swethaa S. A review on ball, roller, low plasticity burnishing process. Materials Today: Proceedings 2019; 18(7): 5087-5099.
23. Pu Y, Zhao Y, Zhang H, Zhao G, Meng J, Song P. Study on the three-dimensional topography of the machined surface in laser-assisted machining of Si3N4 ceramics under different material removal modes. Ceramics International 2019; In press, Available online 14 November 2019.
24. Rao D S, Suresh H H, Komaraiah M, Kempaiah U N. Investigations on the effect of ball burnishing parameters on surface hardness and wear resistance of HSLA dual-phase steels. Materials and Manufacturing Processes 2008; 23: 295-302.

25. Revankar G D, Shetty R, Rao S S, Gaitonde V N. Analysis of surface roughness and hardness in ball burnishing of titanium alloy. Measurement 2014; 58: 256-268.
26. Revankar G D, Shetty R, Rao S S, Gaitonde V N. Wear resistance enhancement of titanium alloy (Ti–6Al–4V) by ball burnishing process. Journal of Materials Research and Technology 2017; 6: 13–32.
27. Sadowski Ł, Czarnecki S, Hoła J. Evaluation of the height 3D roughness parameters of concrete substrate and the adhesion to epoxy resin. International Journal of Adhesion and Adhesives 2016; 67: 3-13.
28. Saldaña-Robles A, Plascencia-Mora H, Aguilera-Gómez E, Saldaña-Robles A, Marquez-Herrera A, Diosdado-De la Peña J A. Influence of ball-burnishing on roughness, hardness and corrosion resistance of AISI 1045 steel. Surface and Coatings Technology 2018; 339: 191-198.
29. Sedlacek M, Podgornik B, Vizintin J. Influence of surface preparation on roughness parameters, friction and wear. Wear 2009; 266: 482–487.
30. Swirad S, Wdowik R. Determining the effect of ball burnishing parameters on surface roughness using the Taguchi method. Procedia Manufacturing 2019; 34: 287-292.
31. Swirad S, Wydrzynski D, Nieslony P, Krolczyk G M. Influence of hydrostatic burnishing strategy on the surface topography of martensitic steel. Measurement 2019; 138: 590-601.
32. Teimouri R, Amini S, Bagheri Bami A. Evaluation of optimized surface properties and residual stress in ultrasonic assisted ball burnishing of AA6061-T6. Measurement 2018; 116: 129-139.
33. Travieso-Rodríguez J A, Jerez-Mesa R, Gómez-Gras G, Llumà-Fuentes J, Casadesús-Farràs O, Madueño-Guerrero M. Hardening effect and fatigue behavior enhancement through ball burnishing on AISI 1038. Journal of Materials Research and Technology 2019, 8(6): 5639-5646.
34. Valiorgue F, Zmelty V, Dumas M, Chomienne V, Verdu C, Lefebvre F, Rech J. Influence of residual stress profile and surface microstructure on fatigue life of a 15-5PH, Procedia Engineering 2018; 213: 623-629.
35. Yilmaz H, Sadeler R. Impact wear behavior of ball burnished 316L stainless steel. Surface and Coatings Technology 2019; 363: 369-378.
36. Zaleski K. The effect of vibratory and rotational shot peening and wear on fatigue life of steel. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2017; 19(1): 102-107.
37. Zhang P, Liu Z. Enhancing surface integrity and corrosion resistance of laser cladded Cr–Ni alloys by hard turning and low plasticity burnishing. Applied Surface Science 2017; 409: 169-178.
38. Żurowski W. Structural factors contributing to increased wear resistance of steel friction couples. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2012; 14(1): 19–23.
39. Żurowski W, Brzózka K, Górka B. Analysis of surface layers and wear products by Mössbauer spectral analysis. Wear 2013; 297(1): 958-965.
40. Żurowski W, Brzózka K, Górka B. Structure of friction products and the surface of tribological system elements. Nukleonika 2013; 58: 99-103.

Dr inż. Krzysztof Więcławski
Dr hab. inż. Jędrzej Mączak, prof. ucz.
Dr inż. Krzysztof Szczurowski

Wydział Samochodów i Maszyn Roboczych, IP
Politechnika Warszawska
Narbutta 84., 02-524 Warszawa, Polska
E-mail: krzysztof.wieclawski@pw.edu.pl

Przebieg prądowy jako źródło informacji o parametrach sterowania wtryskiwaczem paliwowym wtrysku pośredniego

Electric current waveform as a source of information about control parameters indirect injection fuel injector

Słowa kluczowe: *wtryskiwacz, diagnostyka wtryskiwacza*

Keywords: *injector, injector diagnostics, ciśnienie wtrysku, przebieg prądowy.*

Streszczenie:

Artykuł przedstawia wyniki eksperymentów laboratoryjnych polegających na testowaniu wtryskiwaczy paliwowych stosowanych w silnikach spalinowych z wtryskiem pośrednim. Podczas eksperymentów wykonano wiele cykli dawkowania wtryskiwaczy zmieniając parametry sterowania, dzięki czemu opracowano charakterystyki dawkowania i określono wpływ stosowanych parametrów sterowania na wynikowy przepływ paliwa. Jednocześnie rejestrowano przebiegi napięcia i natężenia prądu elektrycznego w cewce wtryskiwacza, dzięki czemu możliwe było powiązanie charakterystyk prądowych z determinantami pracy wtryskiwacza. Wykazano, iż parametry prądowe są precyzyjnym kryterium oceny pracy zaworu elektromagnetycznego, ponieważ dzięki wykonanej przez prąd pracy powstaje przepływ paliwa. Zatem poprzez obserwację zmian prądu płynącego przez cewkę zaworu, można precyzyjnie monitorować prawidłowość procesu otwierania przepływu oraz natężenie prądu, przy którym przepływ się rozpoczął oraz określać wielkości mechaniczne jak dawka i ciśnienie paliwa. Wynikiem badań jest opracowanie charakterystyki wiążącej ciśnienie paliwa z natężeniem prądu w punkcie podnoszenia iglicy, co jest podejściem nowatorskim. Taka charakterystyka może być wykorzystana w diagnostyce i sterowaniu wtryskiwaczy paliwowych oraz wszelkiego rodzaju zaworów elektromagnetycznych.

Abstract:

The article presents the results of laboratory experiments consisting in testing fuel injectors used in indirect injection internal combustion engines. During the experiments, many injector dosing cycles were made, changing the control parameters. As a result, the dosage characteristics were developed, and the effect of the control parameters used on the resulting fuel flow was determined. Simultaneously, voltage and electric current waveforms in the injector coil were recorded, so it was possible to link current characteristics with the determinants of the injector operation. It has been shown that current parameters are a precise criterion for assessing the operation of the solenoid valve, because the work done by the current creates a fuel flow. Thus, by observing the changes in the current flowing through the valve coil, one can precisely monitor the correctness of the process of opening the flow and the current at which the flow began and determine the mechanical quantities such as the dose and pressure of the fuel. The result of the research is the development of a characteristic linking the fuel pressure with the current intensity at the needle lifting point, which is

an innovative approach. Such characteristics can be used in diagnostics and controlling fuel injectors and all kinds of solenoid valves.

1. Wprowadzenie

Artykuł przedstawia wyniki eksperymentów laboratoryjnych, których celem było powiązanie parametrów prądowych zarządzających pracą wtryskiwacza elektromagnetycznego z jego parametrami mechanicznymi. Wykorzystanie związku parametrów mechanicznych z wielkościami prądowymi pozwala, na podstawie łatwo obserwowalnych parametrów prądowych, nadzorować pracę wtryskiwacza oraz precyzyjnie oceniać jakość poszczególnych faz wtrysku, czyli kontrować stan techniczny układu paliwowego i wtryskiwacza podczas jego eksploatacji. Znajomość relacji tych parametrów można wykorzystać nie tylko w diagnostyce wtryskiwacza, ale również w sterowaniu nim. Wynika to z faktu, że dzięki pracy wykonanej przez prąd powstaje przepływ paliwa, doprowadzanego do wtryskiwacza pod odpowiednim ciśnieniem. Prąd elektryczny, przepływający przez cewkę zaworu, powoduje powstanie strumienia magnetycznego a w wyniku jego działania pojawia się siła magnetyczna działająca na iglicę (element odcinający przepływ paliwa). Siła ta, po pokonaniu wszystkich sił oporu przeciwstawiających się tej akcji, umożliwia podniesienie iglicy. Dodatkowo, połączenie elektryczne cewki wtryskiwacza realizuje jednocześnie jej sterowanie i zasilanie. Reasumując, praca wykonana przez siłę magnetyczną, wynikającą z przepływającego prądu, jest czynnikiem wiążącym prąd z wartościami mechanicznymi. Dzięki dokładnemu zdefiniowaniu powyższych relacji, można precyzyjnie określić rzeczywiste początki i końce kolejnych faz procesu wtrysku paliwa, co z kolei można wykorzystać w zarządzaniu pracą silnika. Taka identyfikacja może pozwolić na opracowanie odmiennej strategii w stosunku do dotychczasowej logiki sterowania. Posiadając określane na bieżąco informacje o rzeczywistym początku i czasie trwania kolejnych faz wtrysku, można opracować sterownik silnika wykorzystujący te informacje. Sterownik taki będzie odpowiednio korygował sterowanie wtryskiwaczami w przypadku wykrycia zmian w rozpoczęciu i zakończeniu dawkowania w przeciwieństwie do obecnie stosowanych metod bazujących wyłącznie na pomiarze jakości spalin i korekcie parametrów wtrysku na podstawie zaprogramowanych wcześniej map [1]. Dodatkową weryfikacją procesu sterowania może być wielkość dawki paliwowej, określona na podstawie całkowania przebiegu prądowego, w ramach danego sposobu sterowania przepływem prądu przez cewkę wtryskiwacza. Dokładna znajomość chwilowych parametrów faz wtrysku może być również wykorzystana do precyzyjnej diagnostyki wtryskiwaczy przyczyniając się do wczesnego wykrycia uszkodzeń, mogących wpłynąć na degradację układu oczyszczania spalin a nawet uszkodzenie silnika. Z tego też powodu wczesne wykrywanie uszkodzeń w układzie paliwowym jest niezmiernie istotne.

Działanie wtryskiwacza paliwowego może być oceniane na podstawie różnych wielkości. Weryfikacji może podlegać sygnał sterujący lub wynik sterowania, czyli jakość generowanej strugi [3]. Kształt, kąt i stopień rozpylenia mieszanki świadczą o jakości otrzymanej dawki. Lee i in. [5], przeprowadził analizę dynamicznego zachowania zaworu elektromagnetycznego, badając jednocześnie zjawisko pola elektromagnetycznego i przepływ przez kanał wylotowy [10]. W swojej pracy Harantova [2] przedstawiła projekt oraz analizę układu sterowania i zasilania wtryskiwacza. Analizowane było napięcie zasilania wtryskiwacza, ciśnienia w układzie paliwowym dla określonej szerokości impulsu sterującego (PWM) [9]. Jednocześnie starano się osiągnąć jak najwyższą wydajności i stabilność pracy wtryskiwacza. W pracy [12], przedstawiono metodę weryfikacji wtryskiwaczy paliwowych na podstawie charakterystyk prądowych oraz możliwość korygowania działania wtryskiwacza poprzez stosowanie odpowiednich algorytmów sterowania. Tan i in. [14] przedstawia, jak można dostosować strategię sterowania wtryskiwaczem, żeby uwzględnić zmiany w rezystancji i indukcyjności jego cewki, uwzględniając efekty starzenia. Nikolić i in. [8] omówił procesy zachodzące w systemach dostarczania paliwa, mianowicie wtrysk paliwa, tworzenie mieszanki, spalanie a także emisję spalin. Wyznaczono funkcję K jako powiązanie między prędkością przepływu paliwa i ciśnieniem paliwa w zależności od jego rodzaju [4] [11]. Wahania ciśnienia paliwa mają zasadniczy wpływ na rozpylanie, spalanie i wielkość przepływającego strumienia paliwa oraz opóźnienia w pracy wtryskiwacza [6] [7]. Stępień [13], opisał proces tworzenia się osadów w kanałach paliwowych wtryskiwacza paliwowego, wpływ powstałych osadów na zmiany parametrów diagnostycznych wskazujących na stopień degradacji wtryskiwacza. W pracy [4] przedstawiono nową metodę wykorzystującą przepływomierze Coriolisa (CFM) oraz nową, opatentowaną technikę przetwarzania sygnału, do pomiaru przepływu paliwa. Wykazano, że jest możliwy pomiar prędkości przepływu poszczególnych wtryskiwaczy w silniku w czasie rzeczywistym, co pozwala na dokładną ocenę procesu wtrysku. Merola i in. [7] przedstawił metodę weryfikacji wtrysku paliwa i procesu spalania za pomocą diagnostyki optycznej za pomocą systemu endoskopowego sprzężonego z kamerą CCD zamontowany w kolektorze dolotowym.

Powyżej przytoczone zostało kilka przykładów aktualnych publikacji, przedstawiających metody oceny funkcjonowania wtryskiwacza paliwowego. Jest to końcowy element układu paliwowego i od jego prawidłowego działania, zależy prawidłowa praca silnika. Uszkodzenia wtryskiwacza mogą spowodować degradację układów oczyszczania spalin (reaktora katalitycznego). Z tych powodów prawidłowa diagnostyka i sterowanie wtryskiwaczem paliwowym jest niezmiernie istotnym zagadnieniem eksploatacyjnym. W kolejnych rozdziałach zaprezentowano, nowatorską metodę diagnostyki wtryskiwaczy paliwowych i wszelkiego rodzaju zaworów elektromagnetycznych. Metoda opiera się na identyfikacji wartości natężenia prądu sterującego wtryskiwaczem w punkcie podnoszenia iglicy. Wartość natężenia prądu w tym punkcie, nie jest wartością stałą. Zależy od kilku czynników wynikających z właściwości wtryskiwacza i parametrów sterowania. Precyzja określania wartości prądowych i szczegółowa analiza determinant sterowania wtryskiwaczem, pozwala na wykorzystanie współzależności wartości natężenia prądu, w charakterystycznych punktach przebiegu prądowego dawkującego wtryskiwacza z parametrami mechanicznymi wtryskiwacza. Na podstawie obserwacji charakterystycznych punktów prądowych, można wykryć praktycznie wszystkie uszkodzenia elektryczne i mechaniczne wtryskiwacza, bez potrzeby wykorzystywania diagnostyki poprzez złącze OBD oraz innych metod w tym opierających się na wymontowaniu wtryskiwacza z silnika pojazdu. Metoda diagnostyki przedstawiona w tym artykule, może być wykonywana w trakcie eksploatacji bez demontażu wtryskiwacza a po odpowiednim zaimplementowaniu w sterowniku silnika, uzyska się narzędzie działające automatycznie, zapewniające wczesne wykrywanie uszkodzeń zarówno wtryskiwacza jak i układu paliwowego. W literaturze dotyczącej diagnostyki wtryskiwaczy paliwowych, nie ma metod opartych na zależności przedstawionej w kolejnym rozdziale 3, niniejszego artykułu.

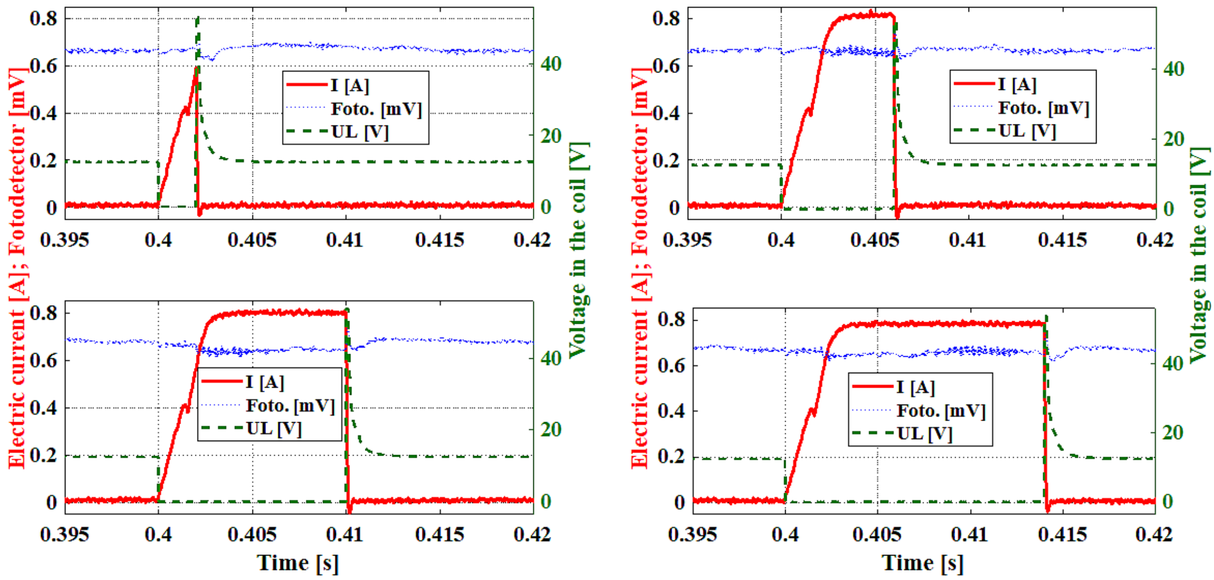
2. Dawka i strumień paliwa

Parametry prądowe precyzyjnie opisują zjawiska, którymi zarządzają. Przedstawia to rys. 1, w którym rosnące pole powierzchni pod przebiegiem natężenia prądu (ciągła, czerwona linia) odzwierciedla coraz większy przepływ określonego medium w odpowiedzi na wydłużanie impulsu napięciowego od 2 ms do 15 ms. Na każdym z przebiegów z rys. 1, na poziomie 0,6 V umieszczono kropkowaną granatową linię określającą wskazania fotodetektora, przetwarzającego natężenie światła lasera na napięcie. W wykonanych badaniach stanowiskowych strumień paliwa zakłócał światło lasera przebiegające pod dyszą wtryskiwacza zmniejszając napięcie fotodetektora o około 0,15 V. Na podstawie tego sygnału można określić rzeczywisty czas przepływu paliwa przez wtryskiwacz.

Praca wtryskiwacza opiera się na prawidłowym współdziałaniu układów elektrycznego i hydraulicznego, sterowanie którymi zapewnia układ elektroniczny. Przepływ prądu w trakcie trwania zadanego czasu wtrysku określa równanie ciągłości:

$$\nabla j + \frac{\partial \rho}{\partial t} = 0 \tag{1}$$

gdzie: j – gęstość prądu elektrycznego,
$\nabla j = \frac{\partial j}{\partial x} + \frac{\partial j}{\partial y} + \frac{\partial j}{\partial z}$ – mnożenie skalarne z operatorem wektorowym nabla,
ρ – gęstość ładunku elektrycznego,
t – czas [s].



Rys. 1. Przebiegi zmian prądu, napięcia oraz natężenia światła lasera odpowiadające zwiększanym czasom wtrysku, od 2 ms do 15 ms zarejestrowane na stanowisku badawczym

Gęstość prądu j, jest wynikiem różniczkowania natężenia prądu względem powierzchni, przez którą przepływa:

$$j = \frac{dI}{dA} \tag{2}$$

gdzie: I – natężenie prądu [A], A – powierzchnia przekroju [m^2].

Ruch iglicy wtryskiwacza wynika ze zmiany strumienia magnetycznego φ, który z kolei wynika z przepływającego przez cewkę prądu I. Mamy wtedy do czynienia z przepływem ładunku elektrycznego Q:

$$Q = \int_{t_1}^{t_2} i(t)dt = \frac{1}{RQ}\int_0^{\varphi} \frac{d\varphi}{dt}dt = \frac{\varphi}{R} \tag{3}$$

gdzie R jest rezystancją [Ω], zaś φ to strumień magnetyczny [Wb].

Dzięki przepływowi ładunków elektrycznych uzyskujemy przepływ paliwa, opisany analogicznym równaniem. Równanie ciągłości strumienia:

$$\rho * \nabla v + \frac{\partial \rho}{\partial t} = 0 \tag{4}$$

We wzorze tym v oznacza objętość $[m^3]$.

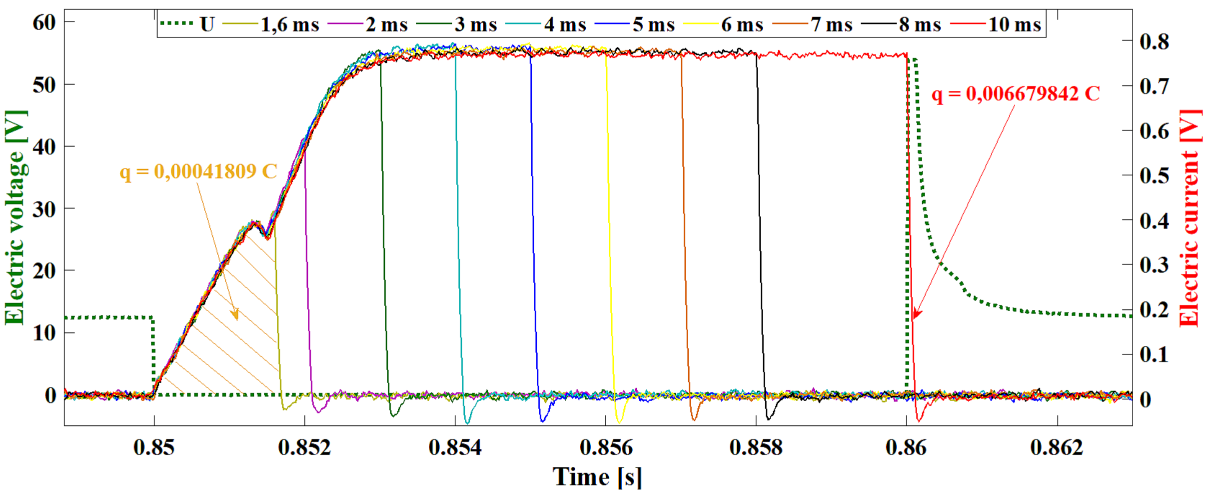
Zmiany strumieni przepływającego paliwa, można odnieść do przepływającego ładunku elektrycznego podczas dawkowania wtryskiwacza. Korzystając z przebiegów prądowych, wykonano charakterystyki dawki oraz strumienia paliwa odpowiadające przebiegom prądowym. Przykładowo, dla czasu wtrysku $10\,ms$ i ciśnienia wtrysku $p = 0{,}3$ MPa, uzyskano przepływ strumienia o wartości: $0{,}0061 \pm 0{,}000124\ \frac{\mathrm{mg}}{\mathrm{ms}}$, co związane jest z przepływem ładunku elektrycznego $Q = 0{,}00668$ C. Ładunek elektryczny Q, obliczono całkując przebieg prądowy (3). Następnie obliczono teoretyczną wielkość strumienia paliwa według równania (5):

$$\dot{m} = f_{s-c} * A * \rho * \sqrt{\frac{2*(p_1 - p_2)}{\rho_i}} \tag{5}$$

gdzie: f_{s-c} – współczynnik przepływu,
ρ_i – gęstość paliwa podczas przepływu przez dyszę wtryskiwacza,
p_1 – ciśnienie paliwa przed wtryskiwaczem,
p_2 – ciśnienie paliwa za wtryskiwaczem.

Dla podanego powyżej przykładu (10 ms *i* 0,3 MPa), otrzymano przepływ masowy o wielkości: $0{,}00556\ \frac{\mathrm{mg}}{\mathrm{ms}}$. Wynik obarczony jest błędem pomiaru, wynikającym z niepewności pomiaru parametrów z równania (5) takich jak: ciśnienie paliwa czy gęstość paliwa. Wynik obliczenia różni się od rzeczywistego o $0{,}000416\ \frac{\mathrm{mg}}{\mathrm{ms}}$, Z uwagi na niepewność pomiaru wielkości z równania (5), jest to wynik zadowalający.

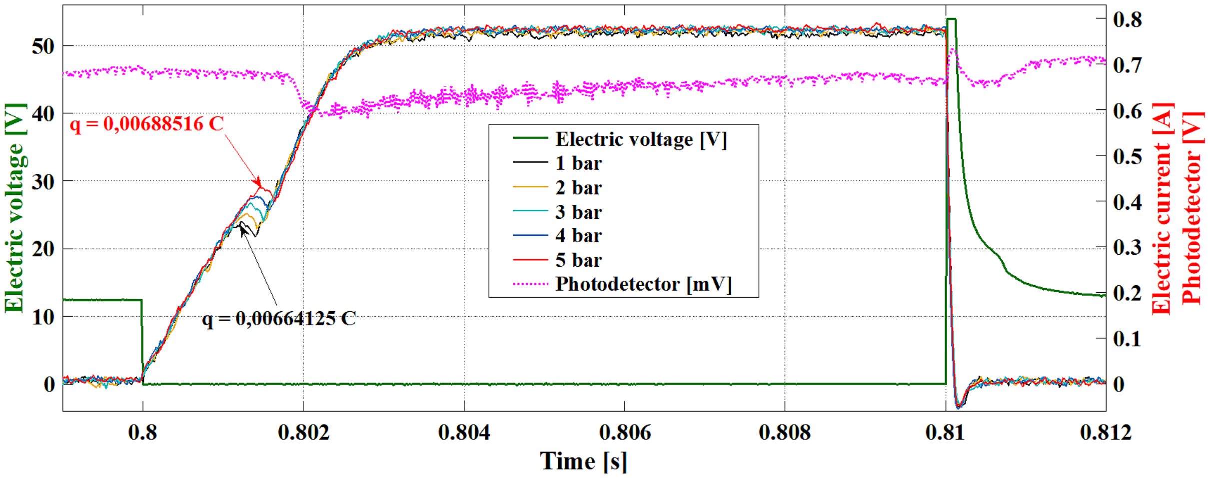
Rys. 2, przedstawia nałożone na siebie przebiegi prądowe dla różnych szerokości impulsów napięciowych, od 1,6 ms do 10 ms.



Rys. 2. Przebiegi prądowe dla różnych szerokości impulsów napięciowych, od 1,6 ms do 10 ms, z przyporządkowaniem pola powierzchni pod przebiegiem.

Do zakresu 1,6 ms (zakres pierwszy z lewej na rys. 2, objęty złotą ciągłą linią), przyporządkowano ładunek elektryczny $q_{1,6ms} = 0{,}000418$ C, natomiast do zakresu 10 ms (cały przebieg prądowy – objęty czerwoną ciągłą linią) ładunek elektryczny o wartości $q_{10ms} = 0{,}00668$ C. Różnica między $q_{1,6ms}$ a q_{10ms} jest 16–krotna. Pola powierzchni kolejnych zakresów jak i ładunki elektryczne są proporcjonalne do odpowiadających im wartości.

W przypadku zakresów prądowych dla różnych ciśnień paliwa (rys. 3), różnice między ładunkami elektrycznymi, reprezentującymi kolejne przebiegi prądowe są zbyt małe, żeby na tej podstawie określić przy jakim ciśnieniu paliwa powstała dawka oraz jaką miała wielkość. Wartości te są w zakresie niepewności pomiaru. Wielkość dawki i strumienia można odnieść do wartości ciśnienia paliwa, określonej na podstawie natężenia prądu w punkcie podnoszenia iglicy, co będzie przedstawione w drugiej części artykułu.



Rys. 3. Przebiegi prądowe dla rosnących ciśnień paliwa od 0,1 do 0,5 MPa.

3. Ciśnienie paliwa

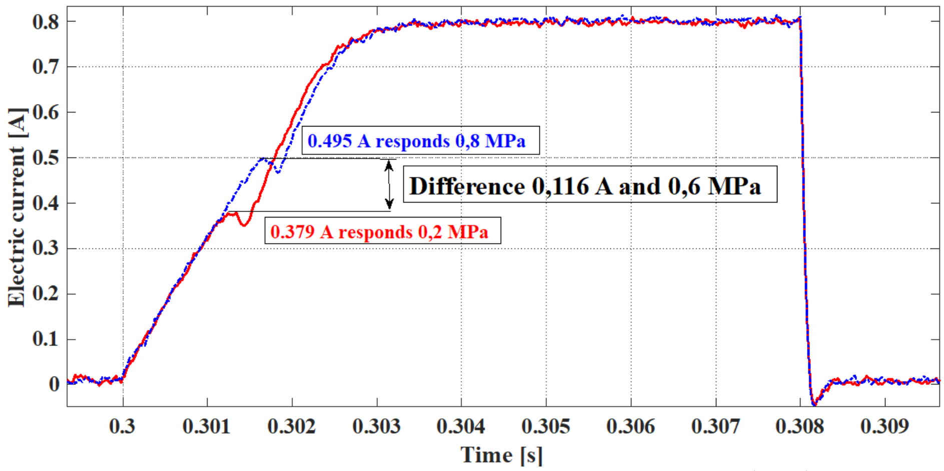
Przebieg prądowy zapisany podczas generowania pojedynczej dawki paliwowej, z zadaną ośmio-milisekundową długością impulsu sterującego przedstawia rys. 4. Obwód elektryczny cewki wtryskiwacza składa się ze źródła siły elektromotorycznej (ε_0), rezystancji (R) i indukcyjności (L) (obwód RL). Wzrost natężenia prądu w obwodzie RL, opisuje równanie Kirchhoffa (6), określając kształt i wartości natężenia elektrycznego zmiennego w czasie.

W przypadku przebiegu prądowego wtryskiwacza, w wyniku wykonywania pracy podnoszenia iglicy przez siłę magnetyczną, równania opisujące natężenie prądu muszą uwzględniać pokonywane przez tę siłę opory, co zostało omówione w tym rozdziale.

$$I_{op} = \frac{\varepsilon_0}{R} * \left(1 - e^{\frac{R}{L}}\right) \qquad (6)$$

Kształt przebiegów czasowych obserwowanych w trakcie pracy zaworu elektromagnetycznego zależy nie tylko od parametrów prądowych (natężenie prądu, napięcie elektryczne) oraz geometrii rdzenia zaworu. Położenie charakterystycznych punktów (zarówno ich wartość jak i chwila czasowa), zależy od gęstości właściwej przepływającego medium oraz ciśnienia z jakim jest ono doprowadzane do zaworu. W tym fragmencie artykułu powołano się na eksperymenty stanowiskowe w których ciśnienie za wtryskiwaczem było stałe. Dlatego w wynikach mowa jest tylko o ciśnieniu przed wtryskiwaczem. Analiza ma za zadanie pokazanie trendu zmian natężenia prądu elektrycznego w punkcie podnoszenia iglicy w zależności od ciśnienia (rzeczywiście, różnicy ciśnień przed i za wtryskiwaczem). Podczas pracy silnikowej wtryskiwacza zmienia się ciśnienie w kolektorze dolotowym, którą to zmianę trzeba uwzględnić w przeprowadzanych analizach.

Wpływ paliwa do kanału wewnętrznego wtryskiwacza, diametralnie zmienia czasowy przebieg napięciowo – prądowy. W wyniku działania gęstości i ciśnienia paliwa (w sumie gęstości, bo ciśnienie powoduje jej wzrost) aby podnieść iglicę należy wygenerować większą siłę magnetyczną. Dla określonej wartości natężenia prądu w punkcie podnoszenia iglicy (I_{op}), można ściśle przyporządkować odpowiadające mu ciśnienie paliwa przed iglicą zaworu (p_{inj}). Rys. 4 przedstawia dwa przebiegi prądowe wtryskiwacza, dla dwóch ciśnień paliwa przed wtryskiwaczem. Zmiana ciśnienia wtrysku z 0,2 MPa do 0,8 MPa, skutkuje zmianą natężenia w punkcie podnoszenia iglicy z 0,379 A do 0,495 A.



Rys. 4. Graficzne przedstawienie zależności $I_{op} = f(p_{inj})$.

Na podstawie badań laboratoryjnych, dla określonego typu wtryskiwacza, można określić funkcję zależności natężenia prądu w punkcie podnoszenia iglicy od ciśnienia wtrysku:

$$I_{op} = f(p_{inj}) \tag{7}$$

Przepływ paliwa uzyskujemy po pokonaniu przez siłę magnetyczną F_m, wszystkich sił przeciwstawiających się podnoszeniu iglicy:

F_p – siły wynikającej z ciśnienia paliwa,

F_i – siły bezwładności,

F_f – siły tarcia,

F_s – siły sprężyny.

$$F_m > F_p + F_i + F_f + F_s \tag{8}$$

Siła magnetyczna F_m jest pochodną energii powstającej w cewce w wyniku działania prądu. Siłę magnetyczną, wystarczającą do podniesienia iglicy, określamy natężeniem prądu I_{op} zmierzonym w tym punkcie przebiegu prądowego wtryskiwacza. Najbardziej znaczącą siłą oporu jest siła F_p, pozostałe składniki są stałe lub zmieniają się z powodu jej zmiany. Z kolei natężenie prądu I_{op}, jest funkcją siły wynikającej z ciśnienia paliwa p_{inj}:

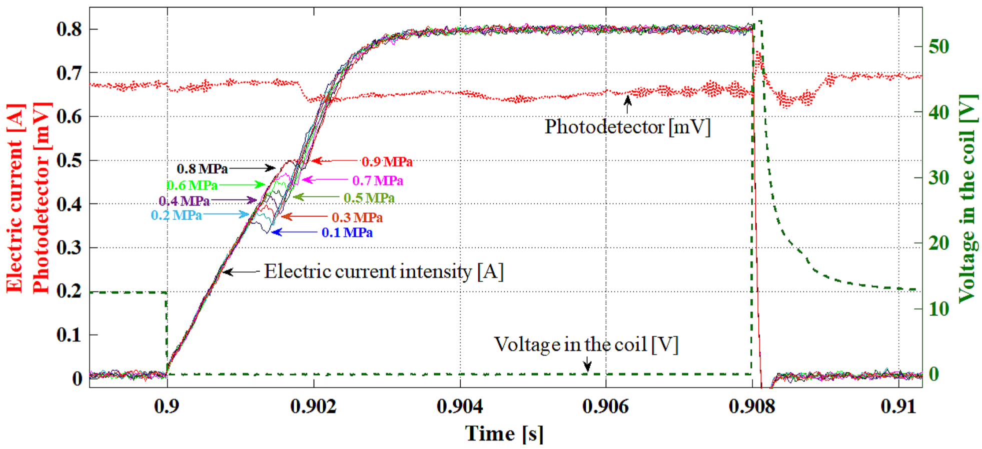
$$F_m = f(I_{op}) \tag{9}$$

gdzie: F_m – siła magnetyczna wynikająca z działania strumienia magnetycznego [N],
I_{op} – natężenie prądu w punkcie podnoszenia iglicy [A],
p_{inj} – ciśnienie wtrysku paliwa [MPa].

Dla wzrastającego ciśnienia paliwa wzrasta siła potrzebna do podniesienia iglicy, zatem wzrasta również prąd potrzebny do wygenerowania strumienia magnetycznego o odpowiedniej wielkości:

$$I_{op1} = f(p_1) < I_{op9} = f(p_9) \qquad \{p_1 < p_9\} \tag{10}$$

Zależność (10) w sposób graficzny przedstawia rys. 5. Pokazano na nim obraz z przebiegu prądowego wtryskiwacza z zaznaczeniem wzrostu natężenia prądu w punkcie otwierania iglicy w zależności od ciśnienia wtrysku.

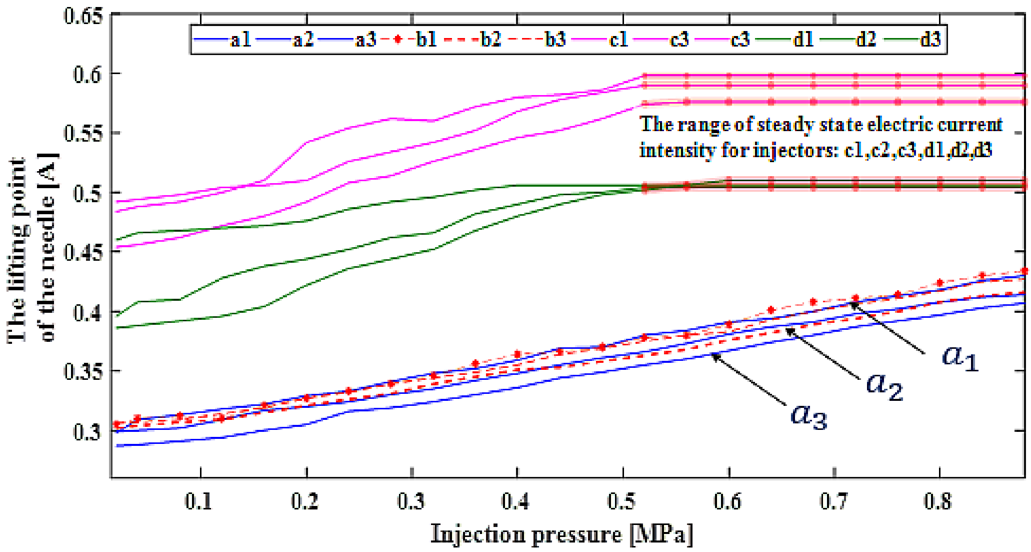


Rys. 5. Zmiany natężenia prądu w punkcie podnoszenia iglicy wtryskiwacza dla różnych ciśnień wtrysku.

Natężenie prądu w punkcie podniesienia iglicy (I_{op}, rys. 5), ściśle odpowiada ciśnieniu wtrysku. Jest ponadto niezależne od czasu wtrysku, a zależy od wielkości generowanej siły magnetycznej. Im większa jest różnica między ciśnieniem przed wtryskiwaczem a ciśnieniem w kolektorze dolotowym (za wtryskiwaczem), tym wyższe jest natężenie prądu I_{op}. Odwzorowanie pracy iglicy, czyli sytuacji, gdy iglica jest przemieszczana, wymaga zastosowania w równaniu (6) dodatkowego współczynnika f_{press}. Po wstawieniu współczynnika w równanie różniczkowe w składnik $\frac{\varepsilon_0}{R}$, otrzymamy oczekiwaną wartość prądu I_{op}:

$$I_{op} = \left(f_{press} * \frac{\varepsilon_0}{R}\right) * \left(1 - e^{\frac{R}{L}}\right) \tag{11}$$

Pierwszy składnik określa wartość maksymalną którą osiągnie I_{op}, a z drugiego składnika wynika eksponencjalne dążenie do tej wartości (równanie Kirchhoffa). Dzięki wyrażeniu (11), można uwzględnić siły oporu w równaniu określającym prąd w obwodzie pracującego wtryskiwacza. Charakterystyka $I_{op} = f(p_{inj})$, konieczna w szczegółowym modelowaniu przebiegu prądowego wtryskiwacza, może służyć do oceny stanu technicznego danego typu wtryskiwacza, w porównaniu różnych rodzajów wtryskiwaczy lub we wczesnym diagnozowaniu jego uszkodzeń. Wszelkiego rodzaju defekty zaworu elektromagnetycznego, będą skutkowały zmianą natężenia prądu w punkcie podnoszenia iglicy. Można monitorować wtryskiwacz w trakcie pracy, kontrolując ten parametr oraz wartość różnicy między ciśnieniem wtrysku i ciśnieniem w kolektorze dolotowym. Charakterystyka $I_{op} = f(p_{inj})$ to zależność o charakterze liniowym. W miarę wzrostu ciśnienia notujemy niewielki wzrost proporcji natężenia prądu I_{op} do p_{inj}. W ramach badań opracowane zostały charakterystyki dla kilku różnych typów wtryskiwaczy z różnym stopniem zużycia. Dla danego typu wtryskiwacza, charakterystyki mają podobny charakter, ale nie są identyczne co wynika z różnic w ich działaniu. Charakterystyki dwóch typów wtryskiwaczy, oznaczonych symbolami „a” i „b” (rys. 6), zmieniają się w całym zakresie ciśnienia wtrysku (od 0,1 do 0,8 MPa).



Rys. 6. Charakterystyki $I_{op} = f(p_{inj})$ dla różnych wtryskiwaczy: “a”, “b”, “c” i “d”.

Strumień magnetyczny wynikający z przepływającego prądu, generowany wokół cewek tych wtryskiwaczy, jest większy niż generowany przez cewki wtryskiwaczy „c” i „d” (rys. 6). Natężenie prądu dla wtryskiwaczy „c” i „d”, począwszy od ciśnienia 0.5 MPa, przechodzi w stan ustalony (maksymalna wartość – pozioma linia). Powyżej tego ciśnienia wtryskiwacze nie zwiększają swoich dawek a natężenie w punkcie podnoszenia iglicy jest stałe (takie jak natężenie prądu stanu ustalonego w obwodzie). Zdolność do pokonania sił przeciwstawiających się podnoszeniu iglicy przy niższym prądzie I_{op}, świadczy o większej sprawności danego wtryskiwacza.

Charakterystyka zależności natężenia prądu w punkcie podnoszenia iglicy od oporów mechanicznych podnoszenia, wynika z właściwości elektrycznych i geometrycznych wtryskiwacza, układu paliwowego i zadanych parametrów sterowania. Dokładne określenie odpowiadających sobie parametrów, pozwala weryfikować stan techniczny układu paliwowego również w czasie bieżącym. To nowatorskie podejście może być wykorzystane zarówno w diagnostyce jak i w korekcie sterowania wtryskiwaczem. Dodatkowo, przedstawiona analiza poparta eksperymentalnie pokazuje, że na podstawie wybranych punktów czasowego przebiegu prądowo – napięciowego, obserwowanego podczas dawkowania wtryskiwacza, można określić wartość różnicy między ciśnieniem wtrysku (ciśnienie przed wtryskiwaczem) a ciśnieniem w kolektorze dolotowym (ciśnienie za wtryskiwaczem) a łącząc te informacje z powierzchnią pola pod wykresem natężenia prądu, można scharakteryzować przepływ masowy.

4. Podsumowanie

W artykule przedstawiono wyniki eksperymentów laboratoryjnych polegających na testowaniu wtryskiwaczy paliwowych stosowanych w silnikach spalinowych z wtryskiem pośrednim bez doładowania. Podczas eksperymentów rejestrowano m.in. przebiegi napięcia, natężenia prądu elektrycznego w cewce wtryskiwacza, opracowano charakterystyki dawkowania oraz określono wpływ stosowanych parametrów sterowania na wynikowy przepływ paliwa. Na tej podstawie wykazano, że parametry prądowe są precyzyjnym kryterium oceny pracy wtryskiwacza. Powiązanie przebiegów prądowych z parametrami dawkowania wtryskiwaczy paliwowych pozwoliło stwierdzić, że zmiany w przebiegach wynikają nie tylko z właściwości prądowych. Zależą również od wartości mechanicznych z których wynika przepływ (4), takich jak gęstość, ciśnienie wtrysku czy opory ruchu iglicy. Wypływa stąd wniosek, że przebieg natężenia prądu elektrycznego (6) w obwodzie cewki wtryskiwacza, zawiera informacje o mechanicznych parametrach wtrysku (11) oraz elektrycznych i mechanicznych właściwości wtryskiwacza. Zakres przebiegu prądowego czy przepływający ładunek elektryczny (3), można przyporządkować do przepływającego strumienia paliwa (5). Dla dokładnego odwzorowania dawki paliwa, konieczne jest określenie przy jakim ciśnieniu przepływ paliwa nastąpił. Charakterystycznym punktem przebiegu prądowego wykorzystanym w przedstawionej analizie, jest punkt podnoszenia iglicy. Wartość natężenia prądu w tym punkcie odnosi się wprost do sił przeciwstawiających się jej podnoszeniu (8). Największą z nich jest siła wynikająca z ciśnienia paliwa, która jest jedyną ze zmiennych. Drugą zmienną jest ciśnienie za wtryskiwaczem, czyli ciśnienie w kolektorze dolotowym. Odniesienie do wartości natężenia prądu w punkcie podnoszenia iglicy ma sens w przypadku takich samych warunków porównania, czyli z bieżącą kontrolą ciśnienia za wtryskiwaczem.

Wynika stąd wniosek, że skokowa zmiana natężenia prądu w tym punkcie następuje w wyniku zmiany stosunku ciśnienia paliwa (7) do ciśnienia w kolektorze dolotowym, co potwierdzają przeprowadzone eksperymenty laboratoryjne.

Obserwacja przebiegów prądowych i ich analiza, w połączeniu z wykonaną wcześniej charakterystyką odpowiadających sobie parametrów, pozwala weryfikować poprawność procesu dawkowania paliwa, jak również służyć ocenie stanu technicznego układu paliwowego i samego wtryskiwacza. Przedstawione przebiegi prądowe mogą być obserwowane przez sterownik w czasie bieżącym w trakcie eksploatacji, dzięki czemu taka weryfikacja może wspomagać diagnostykę OBD oraz systemy kontrolujące skład spalin pozwalając na szybsze wykrycie uszkodzeń w układzie paliwowym. Zależność wartości natężenia prądu w punkcie podnoszenia iglicy od ciśnienia paliwa, może być wykorzystana nie tylko w diagnostyce wtryskiwaczy paliwowych, lecz również w weryfikacji różnego rodzaju zaworów elektromagnetycznych.

Wyniki eksperymentów laboratoryjnych przedstawione w niniejszym artykule, to nowe spojrzenie na diagnostykę wtryskiwaczy paliwowych. Dodatkowo, weryfikacja o której mowa, może być wykonywana w czasie bieżącym w trakcie eksploatacji wtryskiwacza czy zaworu elektromagnetycznego. Diagnostyka „on-line" może przyczynić się do wcześniejszego wykrycia uszkodzeń więc może zabezpieczać przed degradacją układu oczyszczania spalin a nawet uszkodzeniem silnika. Funkcję ciśnienia paliwa można wykorzystać również w sterowaniu wtryskiwaczami. Określenie rzeczywistej fazy wtrysku paliwa może wspomóc zarządzanie pracą silnika. Jest to zmiana strategii w stosunku do dotychczasowej. Przedstawione informacje mogą być wykorzystane w module sterującym pracą silnika, który będzie wykorzystywał w algorytmie sterowania informacje o rzeczywistej chwili rozpoczęcia i zakończenia dawkowania a nie funkcjonował tylko na podstawie opracowanych wcześniej algorytmów, korygowanych przez adaptacje sterowania i sondę lambda.

Acknowledgement: Projekt finansowany przez Narodowe Centrum Badań i Rozwoju w ramach programu DOB-BIO9-04-02-2018.

Bibliografia

1. Ashok B, Ashok D, Ramesh S, Kumar C. A review on control system architecture of a SI engine management system. Elsevier Science. Annual Reviews in Control. 2016, 41; 94-118, doi: 10.1016/j.arcontrol.2016.04.005.

2. Harantová, V., Otáhalová, Z., & Kasanický, M. (2019). Estimation of fuel consumption based on data from opening fuel injector valve. Transportation Research Procedia, 40, 236–243, doi: 10.1016/j.trpro.2019.07.036

3. Kay P J, Bowen P, Bowen P. Gold M R. Sapsford S M. Studies of gasoline direct-injection sprays at elevated ambient gas temperatures and pressures. Atomization and Sprays. 2012; 4 (22), 305–331, doi: 10.1615 / AtomizSpr.2012005411.

4. Leach F, Davy M H, Manus P H, Tombs M, Zhou F. A new method for measuring fuel flow in an individual injection in real time. SAE International Journal of Engines. 2018; 11 (6); 687–696, doi: 10.2307/26649124.

5. Lee G S, Sung H J, Kim H C. (2013). Multiphysics analysis of a linear control solenoid valve. Journal of Fluids Engineering. 2013; 135(1). 011104 – 0111041, doi: 10.1115/1.4023079.

6. Li S, Nehl T, Gopalakrishnan S, Omekanda A, Namuduri C, Prasad R. A Dynamic solenoid model for fuel injectors. IEEE Energy Conversion Congress and Exposition. 2018; 3206 – 3213, doi: 10.1109/ECCE.2018.8558394.

7. Merola, S. S., Tornatore, C., & Sementa, P. (2012). Optical investigation of the fuel injector influence in a PFI spark ignition engine for two-wheel vehicles. Journal of Mechanical Science and Technology, 26(1), doi: 10.1007/s12206-011-0801-5.

8. Nikolić B, Jovanović M, Milošević M, Milanović S. Function K as a link between fuel flow velocity and fuel pressure, depending on the type of fuel. Facta Universitatis, Series: Mechanical Engineering. 2019; 15(1), 119–132, doi: 10.22190/FUME160628003

9. Peng, M., Deng, X., & Liang, X. (2011). Control strategy calibration of fuel injection impulse width on EFI motorcycle engine. 2011 International Conference on Electric Information and Control Engineering, 2047–2053, doi: 10.1109/ICEICE.2011.5776821.

10. Pielecha I, Skowron M, Mazanek A. Evaluation of the injectors operational wear process based on optical fuel spray analysis. Eksploatacja i Niezawodność. 2018; 20 (1): 83–89 doi: 10.17531/ein.2018.1.11.

11. Qiu T, Songa X, Lei Y, Dai H, Cao C, Xu H, Feng X . Effect of back pressure on nozzle inner flow in fuel injector. Fuel. 2016; 173, 79-89, doi: 10.1016/j.fuel.2016.01.044.

12. Sebok, M., Jurcik, J., Gutten, M., Kornciak, D., Roj, J., & Zukowski, P. (2015). Diagnostics and measurement of the gasoline engines injection system. Przegląd Elektrotechniczny, 1(8), doi: 10.15199/48.2015.08.20.

13. Stępień Z. A study of factors influencing the formation of harmful deposits in the diesel engine injectors. Eksploatacja i Niezawodność. 2017; 9 (3), 331–337, doi: 10.17531/ein.2017.3.3.

14. Tan, T., Park, J. S., & Leteinturier, P. (2016). Enhanced Injector Dead Time Compensation by Current Feedback (SAE Technical Paper No. 2016-01–0088). SAE International, doi : 10.4271/2016-01-0088.

dr inż. Joanna Kowalczyk
Department of Mechanical Design
Kielce University of Technology,
Al. 1000-lecia Państwa Polskiego 7, 25-314 Kielce, Poland
jkowalczyk@tu.kielce.pl

dr hab. inż. Monika Madej Prof. PŚK
Department of Mechanical Design
Kielce University of Technology,
Al. 1000-lecia Państwa Polskiego 7, 25-314 Kielce, Poland
mmadej@tu.kielce.pl

prof. dr hab. inż. Dariusz Ozimina
Department of Mechanical Design
Kielce University of Technology,
Al. 1000-lecia Państwa Polskiego 7, 25-314 Kielce, Poland
ozimina@tu.kielce.pl

Ocena właściwości eksploatacyjnych proekologicznej cieczy chłodząco-smarującej zawierającej asparaginian cynku

Słowa kluczowe: *proekologiczna ciecz chłodząco-smarująca, asparaginian cynku, zużycie narzędzia, topografia powierzchni, badania mikrobiologiczne*

Abstract: W pracy przedstawiono wyniki badań wpływu cieczy chłodząco-smarującej z asparaginianem cynku na jakość technologiczną warstwy wierzchniej obrabianych elementów. Asparaginian cynku dotychczas nie był stosowany w takich rozwiązaniach, głównie wykorzystywany był w medycynie i farmakologii. W badaniach przeprowadzono analizę porównawczą proekologicznego chłodziwa zawierającego asparaginian cynku z klasycznym chłodziwem opartym na bazie oleju mineralnego. Ciecze chłodząco-smarujące poddano badaniom toksyczności oraz wykonano kontrolowany proces eksploatacji narzędzi w czasie toczenia poprzecznego. Wyniki badań wskazują, że zastosowanie chłodzenia cieczą na bazie asparaginianu cynku redukuje parametry chropowatości obrabianego materiału nawet o 35%, korzystnie wpływając na jakość finalną detalu.

1. Wstęp

Obróbka skrawaniem jest najczęściej stosowaną metodą kształtowania przedmiotów praktycznie we wszystkich technologiach produkcyjnych [18, 20], w czasie której stosowane są ciecze chłodząco-smarujące [2, 23]. Spełniają one szereg funkcji eksploatacyjnych, spośród których najważniejsze są: chłodzenie, smarowanie oraz poprawa stanu obrabianej powierzchni [22], transportowanie wiórów ze strefy obróbki [3] i czasowa ochrona wyrobu przed korozją [6, 21]. Płyny chłodząco-smarujące, aby mogły spełnić wyżej oczekiwane funkcje muszą mieć określony skład chemiczny, który nie może oddziaływać niekorzystnie na ludzi oraz środowisko [2, 6, 24]. Podczas obróbki skrawaniem stosuje się zazwyczaj ciecze chłodząco-smarujące na bazie wody lub oleju. Substancje smarujące na bazie oleju mineralnego są szeroko stosowane ze względu na niskie koszty. Największą ich wadą jest zanieczyszczanie środowiska, wynikające z trudnej do kontroli degradacji i toksyczności [18, 20]. Dodatkowo płyny te wzbogacane są dodatkami funkcyjnymi, takimi jak modyfikatory tarcia, wysokociśnieniowe EP, przeciwzużyciowe AW, inhibitory korozji, przeciwutleniacze, itp. [1, 3, 10, 16].

Yadav i in. [21] zbadali wpływ dodatków wysokociśnieniowych EP oraz przeciwzużyciowych AW w nowych i zużytych olejach silnikowych klasy SAE 15W40 oraz SAE 20W50 na zużycie badanych kulek. Testy przeprowadzono na czterokulowym testerze Ducom TR-30L, badania realizowano przy następujących parametrach: temperaturze 75^{o}C, obciążeniu 392 N i przy stałej prędkości 1250 obr/min. Uzyskane wyniki wskazały, że średnica wytarcia po testach wzrasta stopniowo wraz ze zwiększaniem czasu eksploatacji olejów silnikowych. Stwierdzono, że działanie dodatków EP i AW determinowane jest przede wszystkim warunkami eksploatacji olejów silnikowych.

Maruda i in. [9] przeprowadzili badania wpływu dodatków wysokociśnieniowych i przeciwzużyciowych (EP/AW) na topografię powierzchni stali podczas toczenia z wykorzystaniem różnych środków chłodzących w zmiennych warunkach prędkości przepływu. Wyniki wykazały, że dodanie dodatków na bazie estru fosforanowego do ośrodka aktywnego spowodowało tworzenie się tribofilmu na styku narzędzie/wiór, który wpłynął na zmniejszenie tarcia. Wykorzystanie metody smarowania minimalną ilością chłodziwa z dodatkami EP/AW doprowadziło do poprawy parametrów topografii powierzchni.

W innym artykule Maruda i in. [10] porównali toczenie: na sucho z chłodzeniem sprężonym powietrzem, mgłą emulsyjną, mgłą emulsyjną + Crodafos O4A-LQ- (MH) i mgłą emulsyjną + Crodafos EHA-LQ- (MH). Wyniki badań wskazały, że użycie podczas toczenia mgły emulsyjnej z modyfikatorami obniżają parametry chropowatości obrabianego materiału nawet o 80%.

Ponadto Maruda i in. [8] zbadali możliwość wprowadzenia dodatków AW i AS we mgle emulsyjnej podczas smarowania najpierw na obrabianej powierzchni elementu maszyny, a następnie do strefy kontaktu pary ciernej. W wyniku tak przeprowadzonych testów uzyskano poprawę jakości powierzchni obrabianej – uzyskano mniejszą chropowatość obrabianej powierzchni stali nierdzewnej, dodatkowo wprowadzone do emulsji dodatki na bazie estrów fosforanowych intensyfikowały to działanie. Wykazano, że dodatek na bazie fosforanów pozostaje na obrabianej powierzchni w znacznych ilościach, nawet po 30 minutach pracy przy dużych obciążeniach, zmniejszając chwilowo współczynnik tarcia i temperaturę w strefie kontaktu.

Na rynek wprowadzane są środki smarowe na bazie oleju roślinnego ze względu na ich biodegradowalność oraz nietoksyczność. Stanowią one potencjalną alternatywę do przygotowania smarów eliminujących pochodne ropy naftowej [15, 17].

Wiele prac poświęcono badaniom cieczy obróbkowych opartych na olejach roślinnych. Ozcelik i współautorzy [14] badali chłodziwa na bazie oleju słonecznikowego oraz rzepakowego z dodatkiem wysokociśnieniowym EP o zawartości 8% i 12%. Wyniki tych badań dowodzą, że ciecz obróbkowa zawierająca olej rzepakowy wykazała lepszą jakość powierzchni obrobionego detalu w porównaniu z olejem słonecznikowym o ponad 11% i 4%, odpowiednio dla zawartości 8% dodatku EP oraz 12% dodatku EP. Kumar i jego zespół [5] badali chłodziwo z olejem kokosowym z dodatkami wysokociśnieniowymi EP. Wykazali, że olej kokosowy zmniejszył siłę posuwu o 31%, siłę nacisku o 28%, siłę skrawania o 20%, temperaturę narzędzia skrawającego o 7% oraz zużycie powierzchni przyłożenia narzędzia o 34% w porównaniu do innych cieczy obróbkowych. Z kolei Zhang i in. [24, 25] badaniom poddali chłodziwa na bazie soi, którą porównano do ropopochodnych cieczy obróbkowych oraz obróbką na sucho. Wykazali, że chłodziwo na bazie soi działało podobnie, jak ropopochodna ciecz obróbkowa. Jednocześnie uzyskano znacznie lepszą chropowatość powierzchni detalu i mniejsze zużycie narzędzia niż po obróbce bez chłodziwa.

Trajano oraz współautorzy [19] przeprowadzili badania tribologiczne oleju słonecznikowego i sojowego zawierającego dodatki CuO oraz nanocząsteczki ZnO. W przypadku nanopłynów sojowych o stężeniu 0,5% wag. CuO i ZnO współczynnik tarcia zmalał odpowiednio o 11% oraz 18%. Natomiast dla oleju słonecznikowego z tą samą frakcją wypełniacza współczynnik tarcia zmalał o 22% i 20%, odpowiednio dla CuO i ZnO.

Z przeprowadzonej powyżej analizy literatury stwierdzono, że w obszarze badań chłodziw istnieją duże luki badawcze. Dotychczas nie prowadzono badań w celu określenia wpływu asparaginianu cynku jako biodegradowalnego dodatku modyfikującego do cieczy chłodząco-smarujących. Analiza bibliometryczna wskazała na zastosowanie asparaginianu cynku jedynie w medycynie i farmakologii. Wykorzystanie asparaginianu cynku jako dodatku do cieczy obróbkowych powinno wpłynąć na poprawę właściwości smarowych oraz utrzymanie w odpowiednim stanie eksploatacyjnym maszyn i urządzeń [4].

W niniejszej pracy przedstawiono wyniki badań uzyskane podczas procesu obróbki skrawaniem z zastosowaniem nietoksycznej cieczy chłodząco-smarującej zawierającej asparaginian cynku. Określono również wpływ na wybrane wskaźniki i jakość powierzchni obrabianego detalu.

2. Materiały stosowane do badań

2.1. Ciecze obróbkowe

Do badań użyto nieobciążającej środowisko cieczy chłodząco-smarującej zawierającej asparaginian cynku o stężeniu 5%, opartej na bazie wody demineralizowanej DEMI. W skład chłodziwa wchodzi, m.in.:

- boran alkanoloaminy;
- biodegradowalny oligomer na bazie poli (kwasu-asparaginowego) (PKA);
- woda demineralizowana.

Podstawą biostabilności cieczy chłodząco-smarującej są poliasparaginiany cynku. Właściwości fizyczne i chemiczne zastosowanej cieczy chłodząco-smarującej zawierającej asparaginian cynku przedstawiono w tabeli 1.

Tabela 1. Właściwości cieczy chłodząco-smarującej zawierającej asparaginian cynku

Kolor	Zapach	*pH*, 3%	Gęstość, g/cm^3	Rozpuszczalność w wodzie
od pomarańczowej do czerwonej	specyficzny	9,2 ÷ 9,7	1,20 ÷ 1,25	rozpuszczalna

Chłodziwo zawierające asparaginian cynku porównano z komercyjną cieczą obróbkową. W jej skład wchodzi głównie olej mineralny, a ponadto: alkohole o długości łańcucha alifatycznego C16÷18, etoksylowane, kwas borowy, dicykloheksyloaminę, 3-lodo-2-propynylbutylkarbaminian, 1,2-benzoizotiazol-3(2H)-jeden. Stosowana jest od ogólnej do ciężkiej obróbki skrawaniem aluminium, stali, żeliwa, metali nieżelaznych, stopów aluminium, mosiądzu oraz miedzi. Chłodziwo to przyczynia się do powstawania dobrej jakości obrabianych powierzchni. Podstawowe parametry cieczy chłodząco-smarującej zamieszczono w tabeli 2.

Tabela 2. Podstawowe parametry porównywanej cieczy chłodząco-smarującej na bazie oleju mineralnego

Kolor	Zapach	Zawartość oleju mineralnego	*pH*, 5%	Gęstość, g/cm^3	Rozpuszczalność w wodzie
żółto-brązowa	jak olej mineralny	56%	9,1	0,92 ÷ 0,96	rozpuszczalna

2.2. Narzędzie i materiał obrabiany

Do toczenia powierzchni czołowych wykorzystano narzędzie wykonane ze stali HS6-5-2C. Stal ta charakteryzuje się bardzo dobrą ciągliwością, udarnością oraz odpornością na ścieranie. Przeznaczona jest ona do pracy na gorąco i może być poddawana obróbce cieplnej w podwyższonych temperaturach: hartowania – 1190 ÷ 1230°C oraz odpuszczania – 550 ÷ 650°C. Jej twardość po ulepszaniu cieplnym w temperaturach 500 ÷ 550°C wynosi 65 HRC. Skład chemiczny stali HS6-5-2C zestawiono w tabeli 3.

Tabela 3. Skład chemiczny stali wolframowo-molibdenowej HS6-5-2C

Pierwiastek	C	Mn	Si	P	S	Cr	Ni	Mo	W	V	Co	Cu
Udział, %	0,82÷ 0,92	max 0,4	max 0,5	max 0,03	max 0,03	3,5÷ 4,5	max 0,4	4,5÷ 5,5	6 ÷ 7	1,7÷ 2,1	max 0,5	max 0,3

Wybór narzędzi ze stali HS6-5-2C był podyktowany tym, że narzędzia te są powszechnie stosowane w obróbce skrawaniem. Szacuje się, że ich udział w produkcji wynosi ok. 40%. Mimo zwiększania się popytu na narzędzia z węglików spiekanych, które można stosować przy dużych prędkościach w obróbce skrawaniem na obrabiarkach CNC, narzędzia ze stali szybkotnących są nadal ekonomiczną alternatywą, biorąc pod uwagę ich cenę oraz łatwość ostrzenia. Coraz częściej na narzędzia wykorzystywane są stale szybkotnące otrzymywane metodą spiekania proszków, a na ich ostrza nanoszone są złożone powłoki dedykowane do konkretnych zastosowań.

Materiałem obrabianym był wałek o średnicy 38 mm wykonany ze stali C45. Jest to stal niestopowa, do ulepszania cieplnego, trudno spawalna, łatwa w obróbce, której skład chemiczny przedstawiono w tabeli 4. Stal C45 stosowana jest na średnio obciążone elementy maszyn i urządzeń. Wyroby z tego materiału mogą być hartowane powierzchniowo, uzyskując twardość 50 ÷ 60 HRC.

Tabela 4. Skład chemiczny stali C45

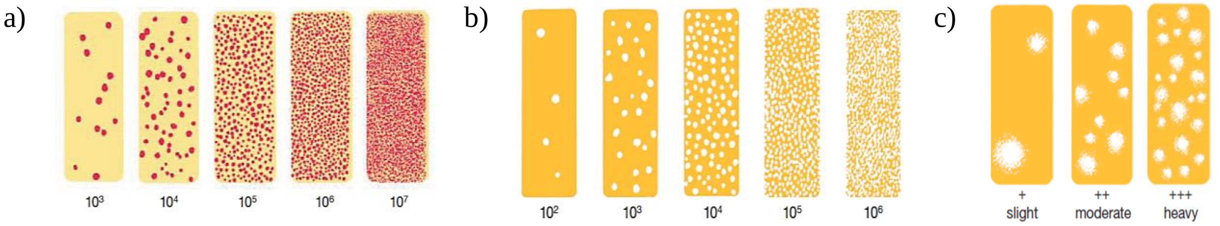
Pierwiastek	C	Mn	Si	P	S	Cu	Cr	Ni	Mo	Cu
Udział, %	0,42÷ 0,5	0,5÷ 0,8	0,1÷ 0,4	max 0,04	max 0,04	max 0,3	max 0,3	max 0,3	max 0,1	max 0,3

3. Metodyka badań

Do badań kinetyki termooksydacyjnej płynów eksploatacyjnych zastosowano spektrometr podczerwieni FTIR Spectrum Two z przystawką ATR firmy Perkin Elmer. Wykonano pomiary próbek czystego oligomeru oraz po przeprowadzonych modelowych badaniach tribologicznych. W czasie badań spektralnych stosowano następujące parametry analityczne:

- zakres spektralny: 4000 ÷ 400 cm^{-1};
- liczba skanów tła: powietrze oraz próbki – 4.

Do oceny mikrobiologicznej – toksyczności chłodziwa użyto specjalny zestaw Microbiology Cult Dip Combi, umożliwiający ustalenie obecności drobnoustrojów. Składał się on z pojemników z przytwierdzonymi płytkami testowymi pokrytymi odpowiednimi pożywkami. Jedna ze stron płytek testowych była jaśniejsza i służyła do wykrywania obecności bakterii, a druga strona – ciemniejsza do wykrywania obecności drożdży i grzybów. Wszystkie części zestawu do badań były sterylne. Testom toksyczności zgodnie z instrukcją [12] poddano chłodziwo zawierające asparaginian cynku oraz konwencjonalne chłodziwo na bazie oleju mineralnego. Pierwsze sprawdzenie próbek nastąpiło po 48 godzinach, drugie po 96 godzinach, a trzecie – po 7 dniach. Po upływie 7 dni próbki poddano ocenie organoleptycznej i porównano ze wzorcami kolonii mikroorganizmów (Rys. 1).



Rys. 1. Wzorce kolonii mikroorganizmów: a) bakterii, b) drożdży, c) pleśni [12]

Proces toczenia poprzecznego przeprowadzono na tokarce sterowanej numerycznie CTX 310 ECO produkcji DMG MORI z wykorzystaniem sterowania Sinumerik 810. Zastosowanie stałej prędkości obrotowej 400 m/min podczas toczenia poprzecznego, spowodowało, że prędkość skrawania v_c dla każdego przejścia zmieniała się cyklicznie w zakresie 47,5 ÷ 0 m/min. W celu porównania właściwości cieczy chłodząco-smarującej zawierającej asparaginian cynku z chłodziwem na bazie oleju mineralnego wykonano toczenie poprzeczne przy użyciu obu cieczy. Po procesie toczenia badano zużycie oraz obserwowano narosty na narzędziach. Parametry toczenia przedstawiono w tabeli 5.

Tabela 5. Parametry skrawania

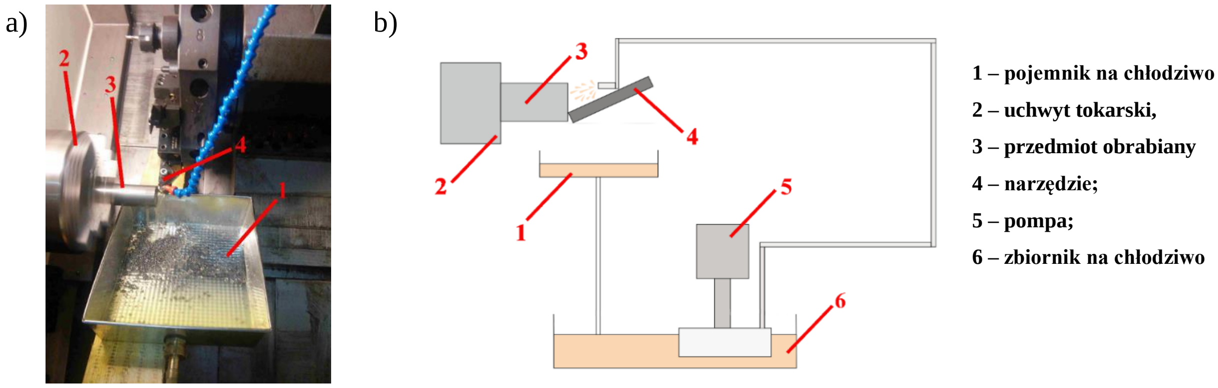
Prędkość skrawania, v_c, m/min	Posuw na obrót, f, mm/obr	Głębokość skrawania, a_p, mm
47,5 ÷ 0	0,098	0,5

Do badań zastosowano wymienne narzędzia skrawające – noże oprawkowe „stalki” o przekroju kwadratowym o wymiarach 10 mm x 10 mm, które mocowano w oprawce do stalek. Stalki wykonane były ze stali szybkotnącej HS6-5-2C. Charakteryzowały się one następującą geometrią:

- kąt przystawienia głównej krawędzi skrawającej K_r = 36,6° w płaszczyźnie podstawowej P_r,
- pomocniczy kąt przystawienia K_r' = 53,4°,
- dodatni kąt natarcia = 5,3°,
- kąt naroża ε_r = 90°,
- kąt podchylenia krawędzi skrawającej λ_s = 7° w płaszczyźnie krawędzi skrawającej P_s,
- kat przyłożenia α = 5,3°,
- kąt natarcia γ = 7°,
- kąt ostrza β = 77,7°,
- promień naroża r_s = 0,04 mm.

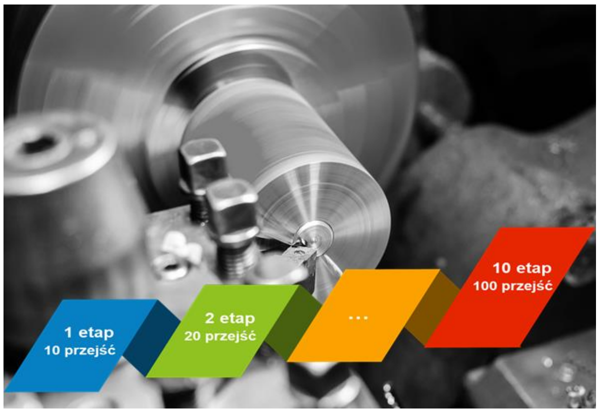
Materiałem obrabianym był wałek ze stali C45 o średnicy 38 mm.

Podczas toczenia zastosowano konwencjonalne chłodzenie zalewowe na powierzchnię natarcia narzędzia. Widok oraz schemat systemu doprowadzania i odprowadzania chłodziwa przedstawiono na Rysunku 2.



Rys. 2. Zewnętrzny system obiegu chłodziwa podczas toczenia poprzecznego na tokarce CTX 310 ECO: a) widok, b) schemat

Wykonano 10 etapów toczenia poprzecznego, z czego pierwszy liczył 10 przejść, a każdy następny był zwiększany o 10 przejść (Rys. 3). Po każdym etapie odcinano cienki kawałek – „plasterek” materiału obrabianego oraz wymieniano narzędzie skrawające.



Rys. 3. Schemat przebiegu procesu toczenia poprzecznego dla noży skrawających

Celem badań była ocena podstawowych funkcji eksploatacyjnych nowej cieczy chłodząco-smarującej zawierającej asparaginian cynku oraz zużycie noża podczas toczenia powierzchni czołowych.

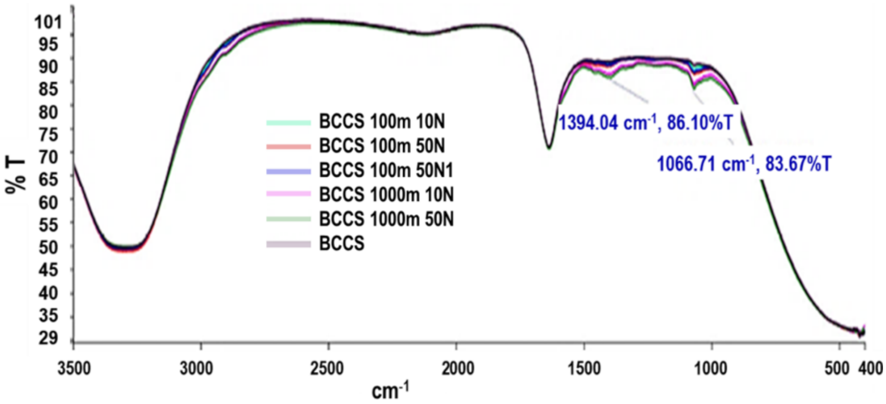
Do oceny topografii powierzchni elementów obrabianych oraz narzędzi skrawających wykorzystano skaningowy mikroskop elektronowy JSM-7100F firmy JEOL. Mikroanalizator EDS umożliwił identyfikację pierwiastków na powierzchni narzędzi skrawających w miejscu powstania narostu, po 10-tym cyklu toczenia z zastosowaniem chłodziw.

Strukturę geometryczną detali po toczeniu obrazowano przy użyciu mikroskopu konfokalnego DCM8 firmy Leica. Dodatkowo do obserwowania zużycia narzędzi skrawających po obróbce wykorzystano stereoskopowy mikroskop inspekcyjny SX80.

4. Wyniki badań

Na Rysunku 4 przedstawiono zestawienie widm, gdzie zaobserwowano wyraźne nasilenie się dwóch pasm wraz ze wzrostem obciążenia i drogi tarcia. Zarejestrowane sygnały o słabej intensywności z maksimum piku przy około 1394 oraz 1066 cm^{-1}, które odpowiadają drganiom rozciągającym wiązania grupy C-O (COH). Świadczy to o zainicjowaniu procesów degradacji chłodziwa i utworzeniem grupy karboksylowej. Obserwacje cieczy chłodząco-smarującej zawierającej asparaginian cynku po testach tribologicznych wskazują zmianę jej barwy. Wraz ze wzrostem obciążenia oraz drogi tarcia staje się ona coraz ciemniejsza, od jasnożółtej do brązowej.

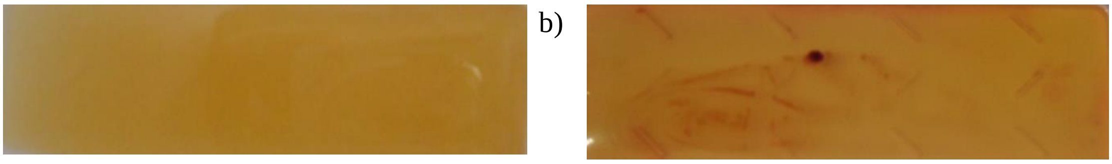
Ciecz chłodząco-smarującą można poddać regeneracji poprzez uzupełnienie składnika podstawowego – asparaginianu cynku oraz kontrolę pozostałych parametrów eksploatacyjnych.



Rys. 5. Zestawienie widm FTIR dla chłodziwa przed i po testach

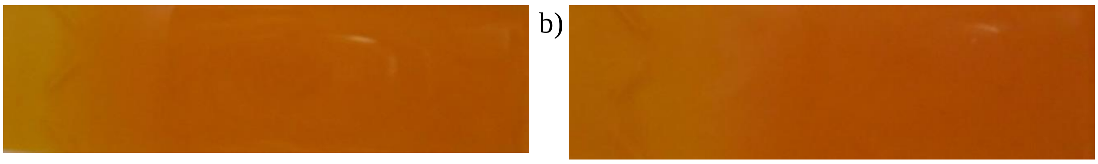
Przy użyciu zestawu Microbiology Cult Dip Combi wykonano ocenę toksyczności chłodziwa zawierającego asparaginian cynku i dla porównania cieczy obróbkowej opartej na bazie oleju mineralnego. Do badań użyto po dwie próbki testowe dla chłodziw poddawanych ocenie. Po upływie 7 dni próbki testowe poddano obserwacjom (Rys. 5 i 6).

a) b)



Rys. 5. Widok próbek testowych po upływie 7 dni – bakterie dla chłodziwa: a) zawierającego asparaginian cynku, b) na bazie oleju mineralnego

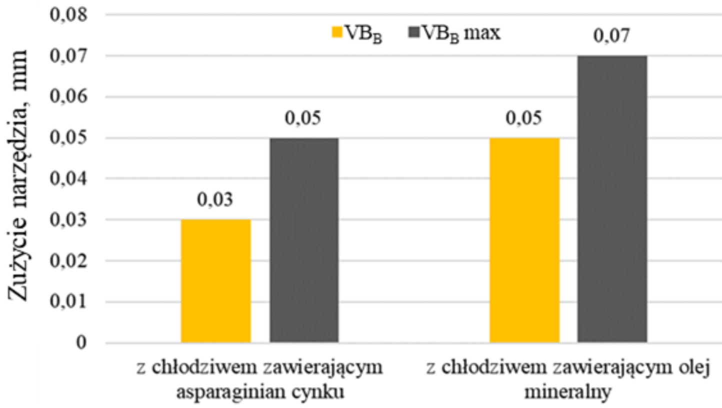
a) b)



Rys. 6. Widok próbek testowych po upływie 7 dni – drożdże i pleśnie dla chłodziwa: a) zawierającego asparaginian cynku, b) na bazie oleju mineralnego

Na podstawie otrzymanych wyników stwierdzono, że ciecz chłodząco-smarująca zawierająca asparaginian cynku jest nietoksyczna. Nie doszło do powstania choćby najmniejszej kolonii bakterii (Rys. 5a), drożdży ani pleśni (Rys. 6a). Natomiast na jednej z próbek testowych, które zanurzone były w chłodziwie na bazie oleju mineralnego zaobserwowano kilka plamek po stronie bakterii (Rys. 5b). Zgodnie z instrukcją [12] powstałe plamki świadczą o powstaniu niewielkiej infekcji. Biorąc pod uwagę uzyskane wyniki chłodziwo zawierające asparaginian cynku okazało się lepsze niż chłodziwo na bazie oleju mineralnego.

Przy użyciu oprogramowania z mikroskopu inspekcyjnego przeprowadzono pomiary zużycia na powierzchni przyłożenia narzędzi skrawających zgodnie z normą [13], które przedstawiono na Rysunku 7.



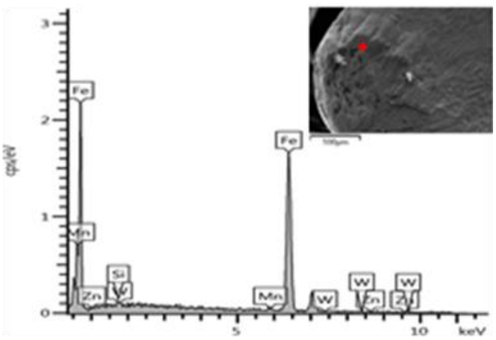
Rys. 7. Wskaźniki zużycia narzędzi skrawających ze stali HS6-5-2C po toczeniu z chłodziwami

Z otrzymanych wykresów wynika, że wartość średniej VB_B oraz maksymalnej VB_Bmax szerokości pasma zużycia osiągnęła niższą wartość równą odpowiednio 0,03 mm i 0,05 mm po obróbce z chłodziwem zawierającym asparaginian cynku. Większe wartości parametrów opisujących zużycie narzędzi otrzymano po toczeniu z zastosowaniem chłodziwa zawierającego olej mineralny.

Na Rysunku 8 zaprezentowano obrazy śladów zużycia ostrzy skrawających po obróbce na sucho i z zastosowaniem chłodziw uzyskane na skaningowym mikroskopie elektronowym oraz wyniki analiz składu chemicznego w mikroobszarach wykonanych metodą EDS. Po ostatniej 10 próbie toczenia z chłodziwami na narzędziach skrawających w miejscu zużycia powstały narosty, których skład chemiczny zbadano za pomocą skaningowego mikroskopu elektronowego.

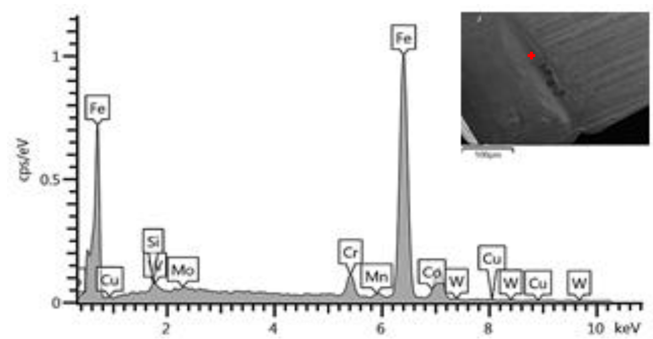
Po toczeniu z cieczą chłodząco-smarującą zawierającą asparaginian cynku w analizowanym punkcie zarejestrowano oprócz pierwiastków wchodzących w skład narzędzia dodatkowo koncentrację atomów cynku. Oznacza to, że powstała cienka warstwa powierzchniowa związków cynku o właściwościach przeciwzużyciowych. Powstała ona jako rezultat procesów tribochemicznych, które zachodzą głównie między dodatkami uszlachetniającymi, zawartymi głównie w środkach smarowych a powierzchniami trącymi. Szybkość i rodzaj reakcji chemicznych jest uzależniony od warunków pracy węzła tarcia. Wytworzona w taki sposób powierzchniowa warstwa reakcyjna zmienia warunki pracy elementów współpracujących tarciowo, co wywołuje kolejne reakcje tribochemiczne. Taka powierzchniowa warstwa wpływa na obniżenie tarcia, a także przedłużenie trwałości eksploatacyjnej elementów współpracujących tarciowo [16].

a)



Zawartość, %							
Si	V	Cr	Mn	Fe	Zn	W	Total
0.76	-	-	1.24	94.43	0.52	0.20	97.15

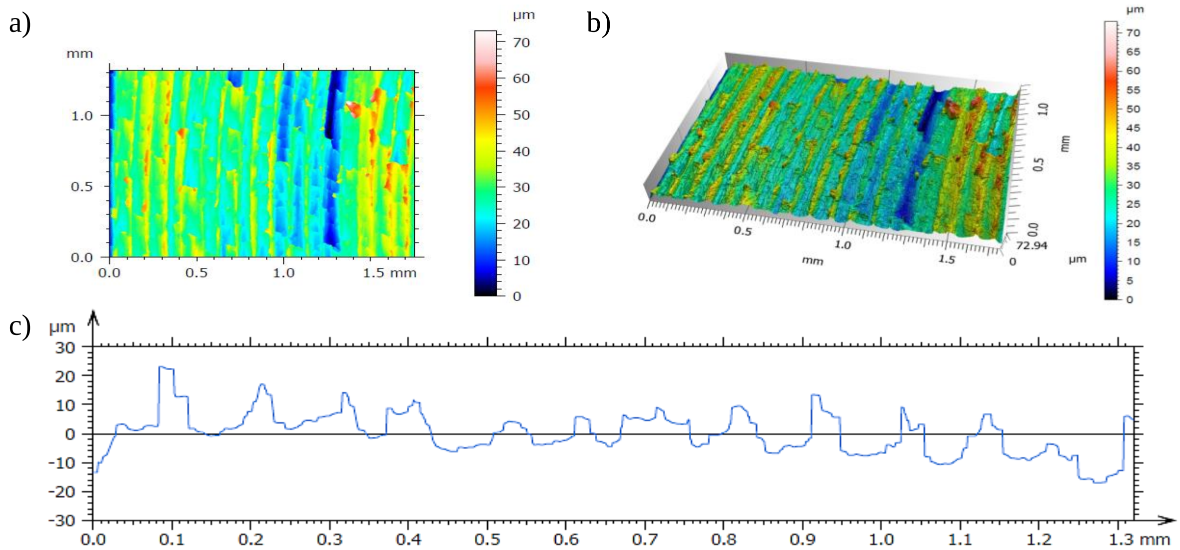
b)



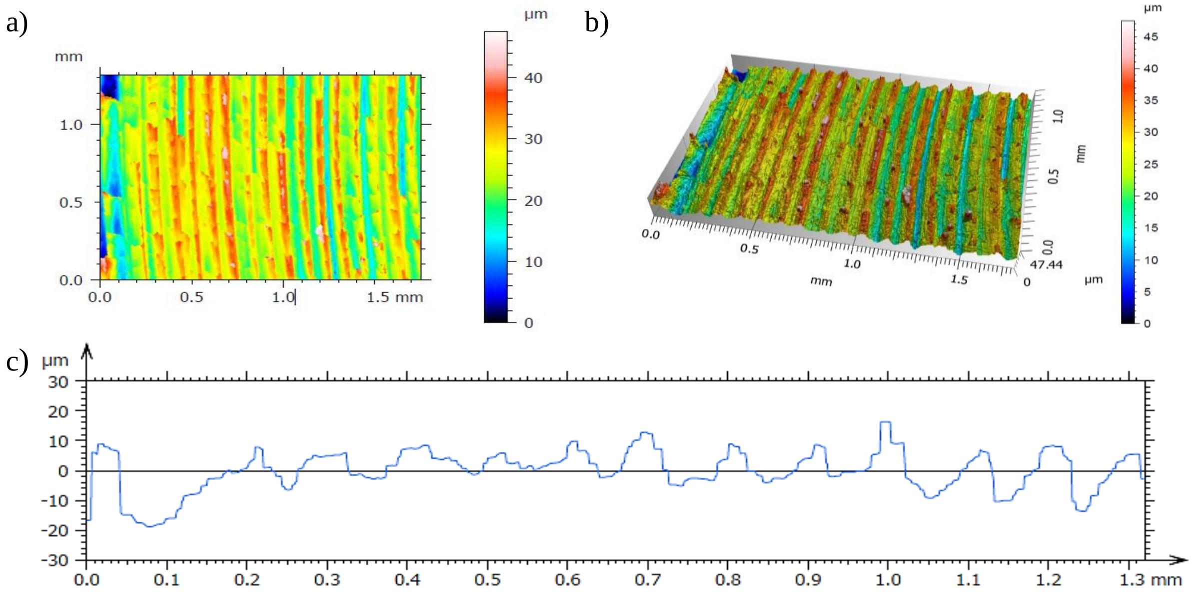
Zawartość, %										
Si	V	Cr	Mn	Fe	Co	Ni	Cu	Mo	W	Total
0.32	-	4.34	0.82	89.12	0.83	-	0.80	0.49	0.65	97.37

Rys. 8. Widmo charakterystycznego promieniowania rentgenowskiego z obszaru śladu zużycia ostrza skrawającego po toczeniu z chłodziwem zawierajacym: a) asparaginian cynku, b) olej mineralny

Na Rysunkach 9 i 10 przedstawiono topografię oraz profile chropowatości powierzchni obrabianych przedmiotów w wybranych obszarach, otrzymanych po 10 etapie toczenia na sucho oraz z cieczami chłodząco-smarującymi.



Rys. 9. Struktura geometryczna brzegu obrobionego przedmiotu po toczeniu z chłodziwem zawierającym olej mineralny: a) mapa warstwicowa, b) widok izometryczny, c) profil pierwotny



Rys. 10. Struktura geometryczna brzegu obrobionego przedmiotu po z chłodziwem zawierającym asparaginian cynku: a) mapa warstwicowa, b) widok izometryczny, c) profil pierwotny

Porównując mapy warstwicowe, widoki izometryczne i profile pierwotne zaobserwowano w obu przypadkach równomierne rozłożenie wierzchołków [7] prawie co 1 mm (posuw wynosił 0,098 mm). Najniższe wzniesienia ok. 16 µm oraz najpłytsze wgłębienia ok. 15 µm powstały na elemencie po toczeniu z zastosowaniem cieczy chłodząco-smarującej zawierającej asparaginian cynku. Natomiast najwyższe wzniesienia i wgłębienia sięgały ok. 20 µm odnotowano po toczeniu z zastosowaniem chłodziwa zawierającego olej mineralny.

W tabeli 6 przedstawiono parametry struktury geometrycznej powierzchni materiałów obrobionych powstałych po toczeniu z zastosowaniem cieczy chłodząco-smarującej zawierającej asparaginian cynku i chłodziwa zawierającego olej mineralny.

Tabela 6. Parametry struktury geometrycznej powierzchni materiałów obrobionych po toczeniu

Warunki toczenia	Parametry struktury geometrycznej powierzchni						
	Sa	*Sq*	*Sp*	*Sv*	*Sz*	*Ssk*	*Sku*
	μm	μm	μm	μm	μm	-	-
z chłodziwem zawierającym olej mineralny	6,41	8,44	44,32	28,62	72,94	0,26	4,06
z chłodziwem zawierającym asparaginian cynku	5,00	6,31	22,12	25,32	47,44	-0,22	3,34

Zestawiając wartości parametrów chropowatości powierzchni obrobionych elementów po procesie toczenia poprzecznego z zastosowaniem chłodziwa zawierającego asparaginian cynku odnotowano niższe wartości niż z użyciem chłodziwa z olejem mineralnym. Świadczy to o tym, że ciecz chłodząco-smarująca wpływa na poprawę jakości powierzchni elementów obrabianych.

4. Wnioski

Na podstawie otrzymanych wyników badań sformułowano następujące wnioski:

1. Zastosowane w procesie toczenia poprzecznego chłodziwo z asparaginianem cynku wpłynęło na zmniejszenie zużycia elementów systemu tribologicznego poprzez tworzenie warstw zawierających związki cynku o dobrych właściwościach przeciwzużyciowych.
2. Najniższe wzniesienia oraz najpłytsze wgłębienia powstały na elemencie po toczeniu z zastosowaniem cieczy chłodząco-smarującej zawierającej asparaginian cynku.
3. Morfologia powierzchni w procesie toczenia zależy od kierunku ruchu oraz wielkości posuwu.
4. Nietoksyczne chłodziwo z asparaginianem cynku w procesie obróbki ulegało reakcyjnej przemianie uzależnionej od warunków eksploatacji.
5. Stwierdzono, że ciecz chłodząco-smarująca zawierająca asparaginian cynku nie wykazuje właściwości toksycznych i może być poddawana prostej regeneracji.
6. Badana ciecz chłodząco-smarująca ze względu na dobre właściwości proekologiczne może być stosowana podczas eksploatacji na konwencjonalnych tokarkach, gdzie bezpośredni kontakt operatora z chłodziwem jest większy niż na tokarkach sterowanych numerycznie.
7. Ciecz obróbkowa z asparaginianem cynku zagwarantowała lepszą jakość technologiczną warstwy wierzchniej obrabianych elementów oraz wydłużyła czas pracy narzędzia ze stali szybkotnącej w porównywaniu do klasycznego chłodziwa. Wartości parametrów chropowatości powierzchni obrobionych elementów były niższe (dla parametru Sa o 22%, Sq o 25%, Sp o połowę, Sv o 11%, Sz o 35%, Ssk ponad 2 razy, a Sku o 17%) niż z użyciem chłodziwa z olejem mineralnym.

References

1. Brinksmeier E, Meyer D, Huesmann-Cordes A G, Herrmann C. Metalworking fluids – Mechanisms and performance. CIRP Annals – Manufacturing Technology 2015; 64(2): 605–628.

2. Jain A, Kumar S, Bajpai V, Wook Park H. Replacement of Hazard Lubricants by Green Coolant in Machining of Ti6Al4V: A 3D FEM Approach. The International Journal of Precision Engineering and Manufacturing 2019; 20 (6): 1027–1035.

3. Kępczak N, Rosik R, Pawłowski W, Sikora M, Witkowski B, Bechciński G, Stachurski W. The dynamics of wear of cutting inserts during turning of non-homogeneous material on the example of polymer concrete. Eksploatacja i Niezawodność – Maintenance and Reliability 2018; 3 (20): 478–483.

4. Kowalczyk J. Ocena właściwości wybranych proekologicznych cieczy chłodząco-smarujących w systemach tribologicznych z powłokami przeciwzużyciowymi. Rozprawa doktorska, Kielce: PŚk Pubisher, 2019.

5. Kumar B S, Padmanabhan G, Krishna P V. Experimental Investigations of Vegetable Oil Based Cutting Fluids with Extreme Pressure Additive in Machining of AISI 1040 Steel. Manufacturing Science and Technology 2015, 3: 1–9.

6. Li K M, Liang S Y. Performance profiling of minimum quantity lubrication in machining. The International Journal of Advanced Manufacturing Technology 2007; 35: 226–233.

7. Maruda R, On the surface roughness of 316L austenitic stainless steel after turning when cooling with MQCL method. Mechanik 2016; 8-9: 1058–1059.

8. Maruda R W, Feldshtein E E, Legutko S, Królczyk G M. Improving the efficiency of running-in for a bronze-stainless steel friction pair. Journal of Friction and Wear 2015; 36 (6): 548–553.

9. Maruda R W, Królczyk G M, Wojciechowski Sz, Żak K, Habrat W, Niesłony P. Effects of extreme pressure and anti-wear additives on surface topography and tool wear during MQCL turning of AISI 1045 steel. Journal of Mechanical Science and Technology 2018; 32 (2): 1585–1591.

10. Maruda R W, Legutko S, Królczyk G M, Hloch S, Michalski M. An influence of active additives on the formation of selected indicators of the condition of the X10CrNi18-8 stainless steel surface layer in MQCL conditions. International Journal of Surface Science and Engineering 2015; 9 (5): 452–465.

11. McNutt J, He Q. Development of biolubricants from vegetable oils via chemical modification. Journal of Industrial Engineering Chemistry 2016; 36: 1–12.

12. Microbiology – 100778 Cult-Dip combi – instrukcja badania toksyczności, test bakteriologiczny.

13. Norma PN-ISO 3685 Badanie trwałości noży tokarskich punktowych.

14. Ozcelik B, Kuram E, Huseyin C M, Demirbas E. Experimental investigations of vegetable based cutting fluids with extreme pressure during turning of AISI 304L. Tribology International 2011, 44: 1864–1871.

15. Park K H, Yang G D, Lee D Y. Tool wear analysis on coated and uncoated carbide tools in inconel machining. International Journal of Precision Engineering and Manufacturing 2015; 16(7): 1639–1645.

16. Płaza S, Margielewski L, Celichowski G. Wstęp do tribologii i tribochemia, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2005.

17. Sartori S, Taccin M, Pavese G, Ghiotti A, Bruschi S. Wear mechanisms of uncoated and coated carbide tools when machining Ti_6Al_4V using LN_2 and cooled N_2. The International Journal of Advanced Manufacturing Technology 2018; 95(1 – 4): 1255–1264.

18. Styp-Rekowski M, Matuszewski M, Oborski I L. Possibilities of some constructional materials cutting by means of water-abrasive jet. Journal of Polish CIMAC 2012; 7 (3): 293–300.

19. Trajano M F, Moura E I F, Ribeiro K S B, Alves S M. Study of oxide nanoparticles as additives for vegetable lubricants. Journal of Materials Research 2014, 17, 1124–1128.

20. Wartacz A, Świć A, Zubrzycki J. The waste of tool blade, and a change of turning parameters in next cuts. Eksploatacja i Niezawodność – Maintenance and Reliability 2004; 4 (24): 58–61.

21. Yadav G, Tiwari S, Jain M L. Tribological analysis of extreme pressure and anti-wear properties of engine lubricating oil using four ball tester. Materials Today: Proceedings 2018; 5: 248–253.

22. Yoon H S, Kim M S, Jang K H, Ahn S H. Future perspectives of sustainable manufacturing and applications based on research databases. The International Journal of Precision Engineering and Manufacturing 2016; 17(9): 1249–1263.

23. Zaleski K. The metod of investigations of lubricating properties of cutting fluids, Eksploatacja i Niezawodność – Maintenance and Reliability 2004; 4 (24): 62–65.

24. Zhang J L, Rao P N. Green/Sustainable Manufacturing–Evaluation of a Soybean-Based Metal Cutting Fluid in Turning Operation. Applied Mechanics and Materials 2013, 392: 925–930.

25. Zhang J, Rao P N, Eckman M. Evaluation of bio-based cutting fluids in using multiple quality characteristics through Taguchi design method. International Journal Modern Engineering 2012, 12: 35–44.

Dr inż. Waldemar Paszkowski

Faculty Organization and Management
Silesian University of Technology
Roosevelta 26 str., 41-800 Zabrze, Poland
E-mail: wpaszkowski@polsl.pl

Modelowanie zjawisk wibroakustycznych z zastosowaniem metody parametryzacji sygnału fonicznego

Słowa kluczowe: *zjawiska wibroakustyczne, parametryzacja sygnału, percepcja hałasu, modelowanie*

Streszczenie: W artykule zaproponowano oryginalny sposób modelowania zjawisk wibroakustycznych eksploatowanych maszyn/urządzeń z zastosowaniem metody parametryzacji sygnału fonicznego. Sposób ten rozszerza dotychczasowe podejście do tego rodzaju badań i polega na uwzględnianiu efektów psychoakustycznych towarzyszących emisji energii wibroakustycznej. Proponowane rozwiązanie opiera się na wyznaczeniu współczynników mel-cepstralnych badanego sygnału i jego klasyfikacji, ze względu na oddziaływanie hałasu. Przedstawiono weryfikację zastosowania metody na przykładzie badań oddziaływania źródeł hałasu drogowego.

1. Wprowadzenie

Źródłem zjawisk zachodzących podczas eksploatacji maszyn i urządzeń są złożone procesy wibroakustyczne. Składają się na nie zróżnicowane oddziaływania drgań, hałasu, dźwięków powietrznych i materiałowych, czy też pulsacji medium w przestrzeniach elementów maszyn. Oddziaływania te w postaci energii wibroakustycznej emitowane są do otoczenia. Efekty zjawisk wibroakustycznych stanowią odpowiedź działania maszyn/urządzeń w stosunku do ich elementów lub otoczenia. Niezależnie od postawionego celu badawczego nośnikiem informacji o zjawiskach wibroakustycznych jest sygnał wibroakustyczny, który podlegać może różnym przekształceniom. Analiza sygnału wibroakustycznego jako procesu przetwarzania danych na informacje użyteczne wykorzystywana jest nie tylko w zadaniach utrzymania ruchu [3, 10], diagnozowania [13], niezawodności [9], oceny stanu eksploatacji obiektów technicznych [14], ale także w zadaniach oddziaływania maszyn/urządzeń na otoczenie. Oddziaływanie to analizowane może być w przestrzeniach otwartych, pomieszczeniach oraz w odniesieniu do organizmu człowieka. Za zjawiska wibroakustyczne uznaje się wszelkie przebiegi drganiowe i akustyczne, które powiązane są ze sobą przyczynowo. Stwierdzono, że w badaniach tych zjawisk uwzględniać się powinno następujące zagadnienia [5]:

- czasowy i przestrzenny rozkład charakterystyki opisującej energię pochodzącą ze źródła,
- odpowiedź układu, wibroakustyczna funkcja przejścia,
- współzależność między źródłami.

Zjawiska wibroakustyczne opisywane są za pomocą podstawowych wielkości fizycznych, takich jak: ciśnienie akustyczne, prędkość, przyspieszenie, przemieszczenie, siła [5]. Z kolei sygnał wibroakustyczny reprezentowany może być za pomocą funkcji jednoargumentowej lub wektora [1].

Stosowane metody analiz sygnałów wibroakustycznych wykorzystują m.in. różne przekształcenia Fouriera dzięki czemu uzyskać można widma amplitudowe, fazowe, bądź energetyczne. Transformata Fouriera poprzez zmianę sygnału z dziedziny czasu do dziedziny częstotliwości pozwala pozyskać cenne informacje, szczególnie gdy występuje duża dynamika zmian parametrów sygnału w czasie. W badaniach wibroakustycznych wymaga się nie tylko przeprowadzenia pomiarów parametrów drgań i hałasu, lecz pozyskania także informacji o badanym zjawisku. Przyjęty cel badań w odniesieniu do obiektu technicznego lub jego otoczenia determinuje ustalenie metodyki pomiarów [5].

2. Modelowanie zjawisk wibroakustycznych oraz percepcji dźwięku - przegląd obecnych rozwiązań

Modelowaniu sygnału wibroakustycznego podlegają zjawiska drganiowe i akustyczne powstające podczas eksploatacji maszyn/urządzeń, które analizowane są łącznie. Do analizy procesów wibroakustycznych w otoczeniu wykorzystuje się modelowanie akustyczne maszyn i urządzeń [5]. Stosuje się w tym celu metody modelowania pola akustycznego maszyn/urządzeń dla identyfikacji źródeł dźwięku. Cechy źródeł dźwięku ocenić można poprzez [5]:

- pole akustyczne wytwarzane przez źródło,
- samo źródło jako emitora energii wibroakustycznej.

Metody modelowania zjawisk wibroakustycznych znajdują szerokie zastosowanie w zadaniach m.in. identyfikacji oraz redukcji źródeł drgań i hałasu, analizy procesów wibroakustycznych [11], oceny stanu technicznego maszyn/urządzeń i ich niezawodności, analizy uszkodzeń elementów maszyn/urządzeń, czy też oceny oddziaływania energii wibroakustycznej na organizm człowieka. Przykładowo, metody modelowania diagnostycznego wykorzystują sygnał wibroakustyczny do badania zużycia lub uszkodzenia elementów obiektu technicznego [5]. Wśród znanych metod modelowania tych zjawisk wyróżnić można m.in. metody: ciśnieniowe, natężeniowe, wzajemnościowe, elementów skończonych, inwersyjne, transformaty Fouriera, krótkoczasowej transformacji Fouriera, transformacji falkowej [1, 5].

W analizie zjawisk wibroakustycznych przetwarzaniu podlegają najczęściej dane i informacje związane bezpośrednio ze źródłem dźwięku, człowiekiem lub środowiskiem, z uwzględnieniem zachodzących interakcji pomiędzy: źródłem dźwięku - człowiekiem - środowiskiem. Jednym z celów poszukiwania związków charakteryzujących maszynę/urządzenie w środowisku tj. źródło hałasu, a parametrami pola akustycznego w jego otoczeniu jest formułowanie modeli emisji dźwięku dla potrzeb oceny jego oddziaływania [23]. W tym zakresie, wyznacza się z sygnału wibroakustycznego odpowiednie cechy fizyczne mierzonych wielkości energetycznych. Na tej podstawie określa się stopień szkodliwości drgań lub hałasu, negatywny wpływ zjawisk wibroakustycznych na organizm człowieka, bądź na elementy maszyn/urządzeń lub otoczenie. Danymi wejściowymi do modelowania zjawisk wibroakustycznych są pozyskiwane informacje akustyczne z ustalonych punktów przestrzeni otoczenia maszyn/urządzeń. Dla potrzeb określenia stopnia narażenia w określonym punkcie przestrzeni powinny wykonane być pomiary wielkości charakteryzujących zagrożenie hałasem [6]. W wyniku pomiarów akustycznych otrzymuje się w danym punkcie odpowiednie wartości wskaźników energetycznych [4, 21]. Prowadzone badania nad słuchowymi skutkami hałasu człowieka potwierdzają, że identyfikacja zagrożenia hałasem jedynie za pomocą wskaźników energetycznych jest ograniczona i niewystarczająca [24, 20]. Pomija się w stosowanym podejściu subiektywne znaczenie wrażeń oddziaływania hałasu. Nie jest uwzględniana interakcja wysokości dźwięku i głośności dynamicznie zmieniających się dźwięków. Stosowane w tym zakresie filtry A-, B , C- oraz D- przybliżają odwrócony kształt krzywych o równej głośności, przy różnych poziomach ciśnienia akustycznego. Ważenie filtrem A stało się najczęściej stosowanym współczynnikiem częstotliwości, chociaż nie jest on rozwiązaniem optymalnym dla wszystkich poziomów ciśnienia akustycznego. Z badań nad percepcją dźwięku wynika, że składowe harmoniczne dźwięku o częstotliwościach zawartych w przedziale 1÷5[kHz] są lepiej słyszalne, niż pozostałe [16]. Ma to znaczenie w interpretacji występowania zjawisk elementarnych i złożonych w niskich oraz wysokich zakresach częstotliwości [15]. W badaniach nad wyznaczeniem subiektywnego odczuwania przez człowieka głośności dla dźwięków ciągłych stosuje się model polegający na wyznaczeniu całkowitego (równoważnego) poziomu dźwięku skorygowanego filtrem korekcyjnym słyszenia (A). Zaznaczyć należy, że charakterystyka częstotliwościowa filtru (A) stanowi pewne przybliżenie odbioru wrażeń akustycznych o niskich poziomach dźwięku. Z analizy charakterystyki progu słyszalności wynika, że czułość słuchu jest największa w zakresie częstotliwości średnich 1000÷4000[Hz] i istotnie spada ona w paśmie niskich i wysokich częstotliwości. Zaznaczyć należy, że charakterystyki progów słyszalności, dyskomfortu oraz bólu istotnie różnią się między sobą w funkcji częstotliwości [12]. Analizując mechanizm powstawania wrażeń akustycznych w stosunku do cech dźwięku to stwierdzić należy, że pomiędzy fizycznymi i subiektywnymi cechami dźwięku nie występuje prosta oraz jednoznaczna relacja. Prowadzone badania nad subiektywnymi cechami dźwięku odnoszą się między innymi do opisu i interpretacji wrażeń słuchowych. Otrzymane wyniki badań nad zjawiskiem hałasu potwierdzają, że istotny wpływ na percepcję hałasu mają m.in. psychoakustyczne aspekty dźwięku, struktura czasowa sygnału, kształtowanie się subiektywnych cech dźwięku w dziedzinie: czasu, częstotliwości, czasowo-częstotliwościowej [8, 17, 18, 19]. Podejmowane badania nad oceną jakości sygnałów fonicznych opierają się na stosowaniu dwóch kategorii metod, [26], tj.:

- metod subiektywnych - polegających na dokonywaniu oceny wrażeń słuchowych przez człowieka,
- metod obiektywnych - polegających na stosowaniu przybliżonych modeli matematycznych w celu uwzględnienia mechanizmów percepcji.

W subiektywnych metodach oceny jakości dźwięku wykorzystuje się odsłuchy sygnału ocenianego dla potrzeb określenia odczuwalnego przez słuchacza stopnia podobieństwa obu sygnałów, stopnia ich różnicy, bądź też poziomu dyskomfortu wywołanego obecnością zakłóceń lub zniekształceń. Metody te z jednej strony pozwalają na wartościowanie otrzymania bezpośredniej odpowiedzi od odbiorcy, z drugiej strony indywidualne czynniki subiektywne wpływają istotnie na pojedynczą ocenę. Wśród metod obiektywnych wyróżnić można badania, tj.:

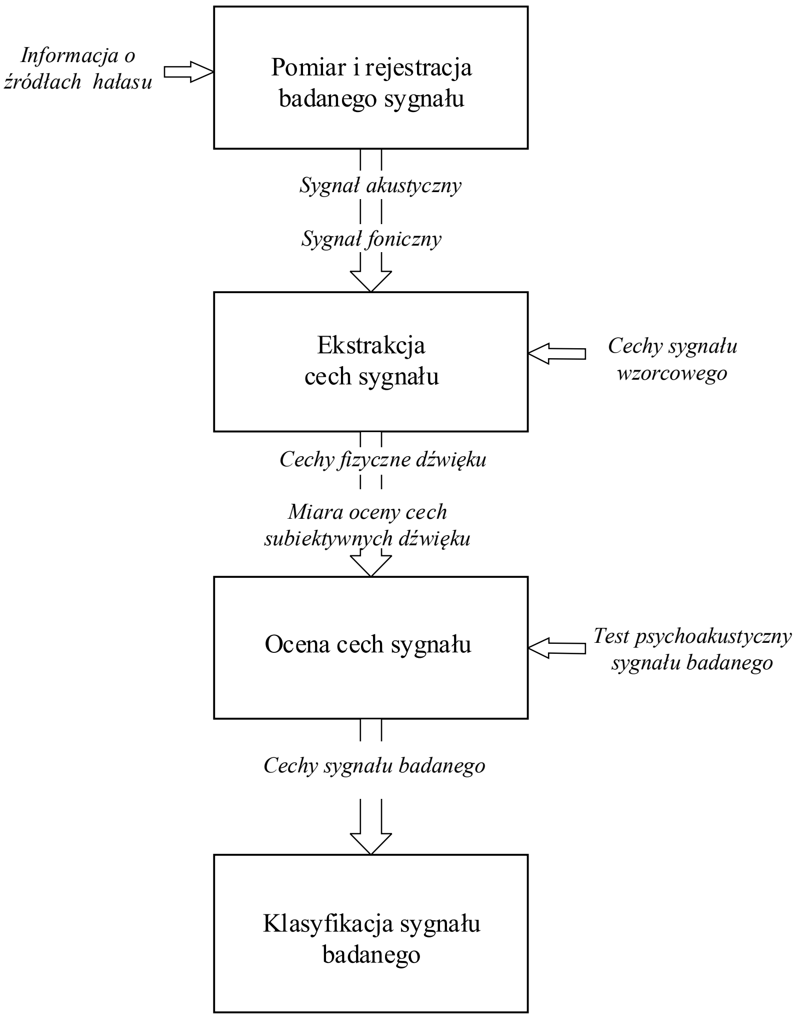
- sygnałowe - najczęściej sygnał oceniany porównywany jest z sygnałem oryginalnym, bez zniekształceń (sygnałem odniesienia). Wykorzystuje się także sposoby, w których nie uwzględnia się sygnału odniesienia.
- parametryczne - ocenie podlega jakość dźwięku na podstawie wiedzy o zastosowanej technice przetwarzania i znajomości jej parametrów, które są argumentami wejściowymi algorytmu oceny.

Z przedstawionej analizy metod i modeli nad stosowanymi rozwiązaniami w zakresie analizy zjawisk wibroakustycznych oraz oceny jakości dźwięku wynika, że cechy zawarte w sygnale fonicznym stanowić mogą cenne źródło informacyjne w badaniu percepcji hałasu. Za ważne w tym obszarze uznać należy uwzględnienie efektów psychoakustycznych oddziaływania energii wibroakustycznej na organizm człowieka zarówno podczas realizacji przez niego prac obsługowo-użytkowych, utrzymania ruchu, jak i przebywania w otoczeniu źródeł emisji drgań i hałasu.

Podejmowany problem dotyczy uwzględnienia w zadaniach modelowania zjawisk wibroakustycznych efektów psychoakustycznych będących skutkiem eksploatacji maszyn/urządzeń.

3. Sposób modelowania zjawisk wibroakustycznych z zastosowaniem metody parametryzacji sygnału

Występujące słuchowe wrażenia dźwiękowe będące skutkiem procesów akustycznych zależne są silnie od częstotliwości, ze względu na uwarunkowania fizyczne propagacji fali akustycznej i percepcję. Za istotne uznać należy nieliniowość i zakres percepcji zjawisk dźwiękowych człowieka, w odniesieniu do amplitudy ciśnienia akustycznego i częstotliwości [2]. Autor zaproponował nowe podejście do modelowania zjawisk wibroakustycznych występujących w otoczeniu eksploatowanych maszyn/urządzeń. Podejście to polega na zastosowaniu ekstrakcji wybranych cech z sygnału akustycznego i sygnału fonicznego. Reprezentację sygnału akustycznego stanowią fizyczne cechy dźwięku. W sygnale fonicznym zawarte są informacje reprezentujące subiektywne cechy dźwięku. W proponowanym sposobie punktem wyjścia do modelowania zjawisk wibroakustycznych są cechy sygnału akustycznego i sygnału fonicznego. Na rys. 1 przedstawiono sposób modelowania zjawisk wibroakustycznych występujących podczas eksploatacji maszyn/urządzeń, który uwzględnia psychoakustyczne efekty hałasu.



Rys. 1 Sposób przetwarzania cech sygnału akustycznego i sygnału fonicznego

Zgodnie z rys. 1 autor zaproponował realizację następujących działań dla potrzeb klasyfikacji sygnału badanego, tj.:

- pomiar i rejestracja badanego sygnału: stanowi źródło informacyjne dla pozyskiwania oraz przetwarzania cech sygnału akustycznego i fonicznego,
- ekstrakcja cech sygnału: polega na zastosowaniu metody/modelu oceny fizycznych i subiektywnych cech sygnału akustycznego oraz fonicznego,
- ocena cech sygnału badanego: podstawą oceny cech sygnału fonicznego jest stopień zgodności jego cech z cechami sygnału wzorcowego,
- klasyfikacja sygnału badanego: przyporządkowanie sygnału fonicznego do danej klasy sygnału wzorcowego na podstawie przeprowadzonej oceny cech.

W przyjętym sposobie założono, że ocena i wybór sygnałów wzorcowych przeprowadzone zostaną na podstawie odtwarzania słuchaczom zarejestrowanych sygnałów fonicznych w ramach eksperymentu psychoakustycznego. Przyporządkowanie sygnałów wzorcowych do odpowiedniej klasy oceny hałasu realizowane jest w oparciu o uzyskane wyniki z prezentowanych sygnałów podczas eksperymentu.

3.1 Opis metody parametryzacji sygnału fonicznego

W rozpatrywanym problemie badawczym zaproponowana metoda parametryzacji sygnału fonicznego opiera się o model perceptualny wykorzystujący własności mechanizmu słyszenia ucha ludzkiego, charakteryzującego się nieliniową percepcją wysokości częstotliwości w skali melowej odbieranych sygnałów dźwiękowych. Zastosowanie przez autora metody parametryzacji sygnału fonicznego w modelowaniu zjawisk wibroakustycznych jest oryginalne i nie spotykane w literaturze przedmiotu. Do najpowszechniej stosowanych metod parametryzacji sygnału mowy należy metoda MFCC (*ang. Mel-Frequency Cepstral Coefficients*), która pozwala na wyznaczenie zbioru współczynników cepstralnych tj. cech sygnału z widma melowego [7]. Współczynniki mel cepstralne wzorowane są na przetwarzaniu sygnału akustycznego w ślimaku narządu słuchu człowieka. Zadaniem ich jest odzwierciedlenie naturalnej odpowiedzi układu słuchowego na pobudzanie dźwiękami. Zaproponowana metoda polega na modelowaniu parametrów ekstrahowanych z sygnału fonicznego, które silnie zależą od subiektywnych wrażeń słuchowych dźwięku. Przesłankami do realizacji badań zjawisk wibroakustycznych za pomocą współczynników mel cepstralnych było rozpoznanie możliwości zastosowania metody parametryzacji sygnału fonicznego, która pozwala na uwzględnienie m.in.:

- losowości sygnałów występujących powszechnie w eksploatowanych maszynach/urządzeniach,
- zmienności struktury częstotliwościowej w przebiegu sygnału,
- estymacji widma sygnału w subiektywnej skali perceptualnej,
- nieliniowości postrzegania częstotliwości dźwięku przez człowieka.

Wyznaczenie współczynników mel-cepstralnych daje możliwości skutecznej klasyfikacji i oceny badanych sygnałów fonicznych. Brak uniwersalności rozwiązań w zakresie rozpoznawania wzorców sygnałów akustycznych dla potrzeb ich oceny nie wynika z niedoskonałości metod, lecz ze złożoności źródłowych sygnałów. Z tego powodu stosuje się transformację analizowanych sygnałów w celu uzyskania odpowiedniej przestrzeni cech dźwięku. Dla potrzeb wyznaczenia wektora współczynników MFCC zastosowano algorytm umożliwiający ekstrakcję cech sygnału fonicznego (rys. 2).

Rys. 2 Procedura parametryzacji sygnału fonicznego z zastosowaniem współczynników MFCC [25]

Zgodnie przedstawioną na rys. 2 procedurą przyjęto etapową realizację przetwarzania cech sygnału [25] tj.:
Etap 1: Proces preemfazy polegający na filtracji formującej, w wyniku której następuje osłabienie składowych o małych częstotliwościach i wzmocnieniu składowych o wysokich częstotliwościach.
Etap 2: Ramkowanie sygnału, czyli podział sygnału na krótkie fragmenty zwane ramkami. Możliwe jest zastosowanie nakładania się kolejnych ramek czasowych. Następnie na tym etapie realizowane jest okienkowanie z zastosowaniem okna Hamminga:

$$Ham(N) = 0,54 - 0,46\cos(2\pi \frac{n-1}{N-1}) \tag{1}$$

gdzie:
N - długość ramki,
n=1,2,...,N.
Etap 3: Wykonanie algorytmu szybkiej transformaty Fouriera (FFT) na zokienkowanym sygnale w poszczególnych ramkach i wyznaczenie modułu estymaty widmowej gęstości mocy sygnału.
Etap 4: Przeprowadzenie filtracji melowej za pomocą zestawu środkowoprzepustowych filtrów trójkątnych o częstotliwościach wyznaczonych zgodnie z:

$$f_{mel} = 2595\log_{10}(1 + \frac{f_{Hz}}{700}) \tag{2}$$

W obliczeniach używany jest logarytm energii, co pozwala na zredukowanie wrażliwości filtrów na bardzo głośne i bardzo ciche dźwięki oraz modelowanie nieliniowej amplitudowej wrażliwości ucha ludzkiego.
Etap 5: Ostatnim etapem procedury jest zastosowanie dyskretnej transformaty kosinusowej (DCT). Otrzymany wektor współczynników MFCC obliczony jest zgodnie z zależnością:

$$MFCC_n = \sqrt{\frac{2}{N}\sum_{i=1}^{N}\log(S_i)\cdot\cos\left[\frac{\pi n}{N}(i-0,5)\right]} \tag{3}$$

$$S_i = \sum_{k=1}^{N} \left|X_r(k)\right|^2 H_i(k) \tag{4}$$

gdzie:
i -numer filtra,
X_r-widmo ramki,
H_i-zestaw filtrów
S_i-energia pasma
n-numer współczynnika,
N-ilość użytych filtrów.
W większości systemów rozpoznawania n przyjmuje wartości od 1, a współczynnik $MFCC_0$ jest pomijany. Wygenerowany wektor współczynników $\mathrm{MFCC_n}$ przyjmuje postać:

$$MFCC_n = \langle MFCC_1, MFCC_2, MFCC_3, ...MFCC_{13} \rangle \tag{5}$$

W przedstawionym sposobie badań założono, że rozpoznane w ramach eksperymentu psychoakustycznego sygnały wzorcowe poddane zostaną parametryzacji celem wyznaczenia współczynników mel-cepstralnych. Zgodnie z procedurą (rys. 2) dla każdego badanego sygnału fonicznego wyznaczane są współczynniki mel-cepstralne, które podlegać będą następnie ocenie przy wykorzystaniu sygnałów wzorcowych. Na podstawie wyznaczonych współczynników MFCC klasyfikacja badanego sygnału polega na ocenie stopnia zgodności jego cech z cechami sygnałów wzorcowych. W celu przeprowadzenia klasyfikacji badanych sygnałów fonicznych (na podstawie zbioru sygnałów wzorcowych) autor zaproponował jako miarę oceny najmniejszą odległość pomiędzy dwoma ciągami wektorów cech, tj. współczynników MFCC z zastosowaniem metody DTW (*ang. Dynamic Time Warping*). Sygnały emitowane przez źródła wibroakustyczne charakteryzuje dynamiczna zmienność w czasie, co powoduje, że cechy tych sygnałów podlegają także zmienności. Dynamiczne dopasowanie czasowe (DTW) należy do metod stosowanych w rozpoznawaniu mowy, w szczególności metoda ta znalazła przede wszystkim zastosowanie w rozpoznawaniu izolowanych słów oraz wyszukiwaniu haseł. Metodę tą wykorzystuje się w rozpoznawaniu i klasyfikacji macierzy cech współczynników MFCC jako nieliniową transformację czasową. Polega ona na transformacji osi czasu, w celu lepszego dopasowania dwóch sekwencji czasowych. Wyznaczenie odległości metodą dynamicznego dopasowania czasowego (DTW) polega na, tj.:

- obliczeniu tzw. macierzy odległości lokalnych *d(m,n)*, która powstaje poprzez obliczenie odległości euklidesowych pomiędzy każdym wektorem badanego sygnału i sygnału wzorca,
- wyznaczeniu sumy odległości lokalnych (euklidesowych), która stanowi odległość zakumulowaną wzdłuż ścieżki optymalnej przebiegającej w macierzy odległości lokalnych, od lewego dolnego rogu macierzy do jej prawego górnego rogu.

Poszczególne odległości DTW pomiędzy sygnałami wzorcowymi i sygnałem badanym wyznaczyć można zgodnie z następującą formułą:

$$d_{mn}(X,Y) = \sqrt{\sum_{k=1}^{K}(x_{k,m} - y_{k,n}) \cdot (x_{k,m} - y_{k,n})} \qquad (6)$$

gdzie:
K- wymiar sygnałów,
m, n - ciągi wektorów cech X i Y.
Odległość zakumulowana obliczona wzdłuż ścieżki optymalnej jest najmniejszą z możliwych odległości zakumulowanych poszczególnych wektorów współczynników MFCC badanego sygnału i sygnału wzorca. Przyjęto, że kryterium klasyfikacji badanego sygnału w zakresie oceny dokuczliwości hałasowej stanowić będzie najmniejsza odległość DTW pomiędzy współczynnikami mel-cepstralnymi badanego sygnału i danym sygnałem wzorcowym o jednakowej długości czasowej.

3.2 Przykład zastosowania metody i klasyfikacja badanego sygnału fonicznego

Wybranym do badań sygnałem fonicznym była próbka dźwiękowa o nazwie DW13a_9_56 charakteryzująca się równoważnym poziomem dźwięku L_{Aeq}=56[dB(A)]. Sygnał zarejestrowany został w punkcie przestrzeni zlokalizowanym na wysokości pasa drogowego i obejmował pomiar imisji źródeł hałasu drogowego przemieszczających się pojazdów po nawierzchni z kostki brukowej. Z uwagi na postawiony cel badań pomiar nie został przeprowadzony zgodnie z obowiązującą metodyką realizacji pomiarów środowiskowych [22]. Do nagrań i pomiarów wykorzystano następujący zestaw urządzeń oraz pomocy pomiarowych:

- miernik poziomu dźwięku Brüel & Kjær 2238 Mediator,
- mikrofon pomiarowy Brüel & Kjær 4188 do pomiarów pola swobodnego,

- rejestrator dźwięku ZOOM Handy Recorder H4n,
- statyw mikrofonowy.

Czas nagrań i pomiarów ustalony został na 5[min]. Podczas nagrywania i pomiarów mikrofon umieszczony był na statywie i w połowie szerokości chodnika, na wysokości 1,7[m], co odpowiadało przybliżonej wysokości położenia głowy potencjalnego przechodnia. Przyjęto, że w ocenie subiektywnej próbek dźwiękowych (w eksperymencie psychoakustycznym) udział będą brać osoby ze słuchem prawidłowym. Dla każdego 5-minutowego pliku wyselekcjonowano dziesięć plików 10-sekundowych. W przeprowadzanym eksperymencie wzięło udział 80 osób w wieku pomiędzy 22-50 lat spełniających ww. założenie. W realizowanych badaniach psychoakustycznych zaproponowano skalę punktową w zakresie od 1 do 5 dla oceny dokuczliwości hałasu drogowego prezentowanych 30 sygnałów fonicznych. Przyporządkowano odpowiednie etykiety ocenianym sygnałom, tj:

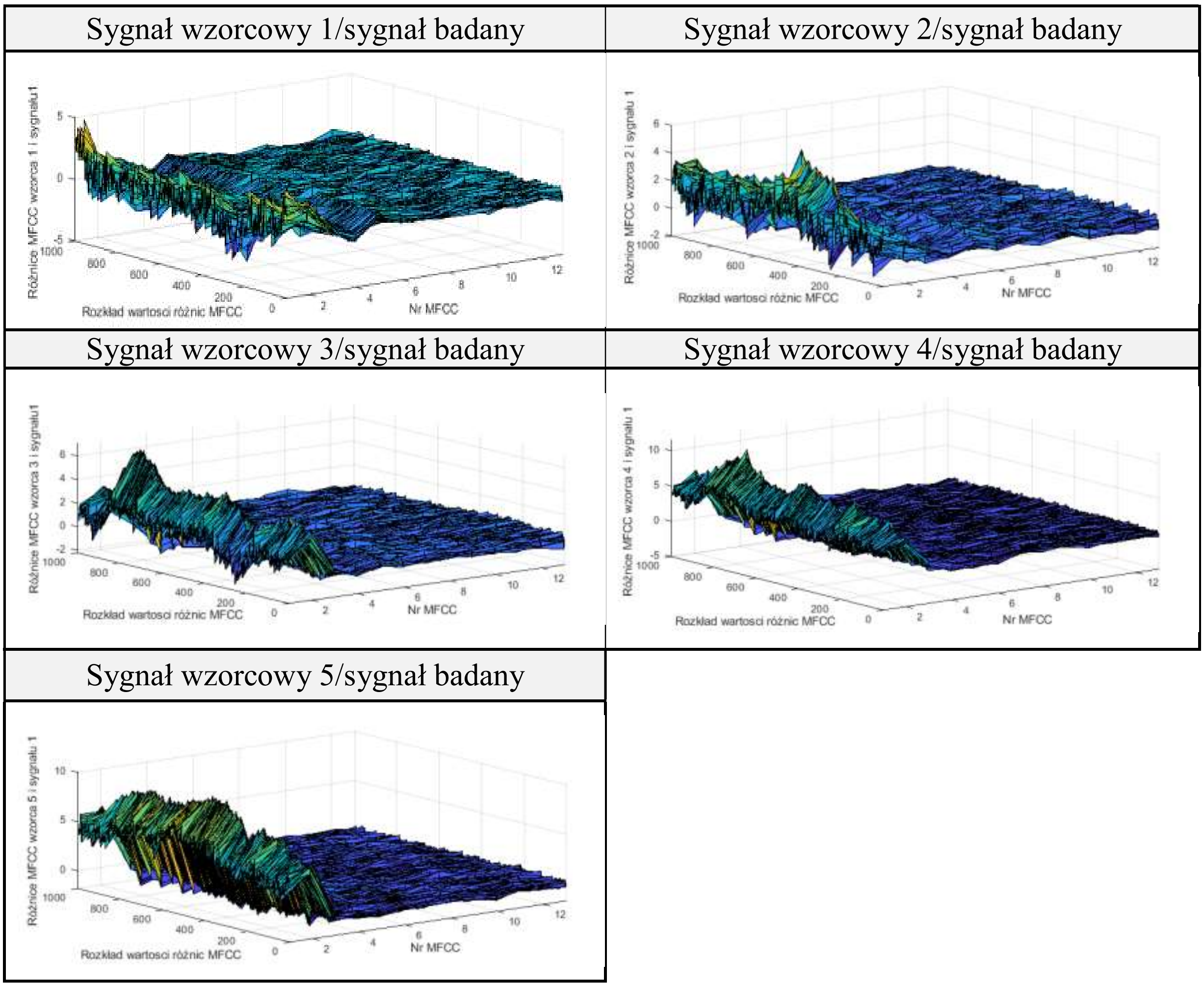
- wcale - ocena 1,
- mało - ocena 2,
- średnio - ocena 3,
- bardzo - ocena 4,
- nie do zniesienia ocena 5.

Biorąc pod uwagę rozpiętość uzyskanych wartości poziomu dźwięku odpowiadającym różnym lokalizacjom pomiarowych (charakteryzującym się zmienną zabudową, zmiennym układem infrastruktury drogowej) oraz różnym nawierzchniom jezdni (tj. bruk, asfalt) zdecydowano, że wyselekcjonowane próbki sygnałów przygotowane zostaną w taki sposób, aby ich poziom równoważny odpowiadał jednej z pięciu następujących wartości, tzn.: 56, 62, 68, 74, 80[dB(A)]. Danymi wejściowymi dla wyznaczenia wektora $MFCC_n$ stanowiły foniczne sygnały wzorcowe w dziedzinie czasu, wybrane na podstawie eksperymentu psychoakustycznego i reprezentujące odpowiednie klasy dokuczliwości hałasowej. W wyniku realizacji procedury (rys. 2) obliczonych zostało 13 współczynników mel-cepstralnych dla każdego ze wzorcowych sygnałów i sygnału badanego. Obliczenia przeprowadzono w środowisku Matlab R2018a przy zastosowaniu funkcji (mfcc). Wartości odchylenia standardowego (σ) oraz średniej arytmetycznej ($\overline{x}$) wyznaczone zostały dla każdego ze współczynników dla 1000 danych (wartości obliczone zostały z krokiem 0,01[s]), (tabela 1).

Tabela 1. Zestawienie wartości σ oraz $\overline{x}$ współczynników mel-cepstralnych $MFCC_n$ dla sygnałów wzorcowych

	Sygnał wzorcowy 1		Sygnał wzorcowy 2		Sygnał wzorcowy 3		Sygnał wzorcowy 4		Sygnał wzorcowy 5	
Nr współcz.	σ	$\overline{x}$	σ	$\overline{x}$	σ	$\overline{x}$	σ	$\overline{x}$	σ	$\overline{x}$
$MFCC_1$	1,25	-2,07	1,40	-2,08	1,16	-1,99	1,38	0,61	1,40	1,43
$MFCC_2$	1,27	-18,01	1,81	-16,32	1,62	-14,21	2,01	-11,14	2,00	-10,60
$MFCC_3$	0,35	3,26	0,31	3,20	0,35	3,11	0,43	3,27	0,43	2,79
$MFCC_4$	0,25	-0,51	0,25	-0,65	0,24	-0,99	0,33	-0,75	0,61	0,13
$MFCC_5$	0,23	0,38	0,23	0,30	0,20	-0,09	0,28	-0,30	0,36	0,03
$MFCC_6$	0,20	0,06	0,22	0,03	0,20	-0,03	0,31	-0,07	0,32	0,03
$MFCC_7$	0,18	0,28	0,18	0,29	0,18	0,16	0,24	0,24	0,29	0,26
$MFCC_8$	0,20	0,11	0,19	0,08	0,18	-0,09	0,21	0,24	0,32	0,19
$MFCC_9$	0,18	0,13	0,20	0,16	0,18	0,13	0,21	0,47	0,34	0,18
$MFCC_{10}$	0,17	0,02	0,18	0,03	0,17	0,10	0,20	0,02	0,37	0,14
$MFCC_{11}$	0,16	0,06	0,17	-0,03	0,18	0,06	0,20	0,18	0,39	0,04
$MFCC_{12}$	0,17	0,02	0,18	-0,02	0,17	-0,01	0,27	-0,22	0,42	-0,02
$MFCC_{13}$	0,16	0,03	0,17	-0,03	0,18	0,02	0,18	0,20	0,44	-0,01

Przeprowadzona analiza porównawcza wartości współczynników MFCC badanego sygnału z sygnałami wzorcowymi wykazała, że największe (lokalne) różnice względne pomiędzy współczynnikami MFCC wystąpiły dla przypadku zestawienia sygnału badanego z sygnałem wzorcowym 4 (rys. 3).



Rys. 3 Analiza porównawcza współczynników MFCC badanego sygnału z sygnałami wzorcowymi

Poszczególne odległości DTW pomiędzy sygnałami wzorcowymi i sygnałem badanym wyznaczone zostały za pomocą funkcji (dtw) w środowisku Matlab R2018a. Po podstawieniu do (6) wyznaczone zostały odległości DTW pomiędzy współczynnikami MFCC dla poszczególnych sygnałów wzorcowych i badanego sygnału fonicznego DW13a_9_56.

Tabela 2. Zestawienie odległości zakumulowanych DTW sygnału badanego z sygnałami wzorcowymi

Lp.	Rodzaj wariantu	Odległość zakumulowana DTW
1.	Sygnał wzorcowy 1 / sygnał badany	197,829
2.	Sygnał wzorcowy 2 / sygnał badany	249,447
3.	Sygnał wzorcowy 3 / sygnał badany	304,864
4.	Sygnał wzorcowy 4 / sygnał badany	496,279
5.	Sygnał wzorcowy 5 / sygnał badany	537,361

Z tabeli 2 wynika, że najmniejszą zakumulowaną odległość DTW uzyskano dla wariantu oceny badanego sygnału DW13a_9_56 z wzorcem 1, co oznacza, że sygnał badany zakwalifikowany został do klasy dokuczliwości hałasu: wcale.

4. Dyskusja wyników i wnioski

Otrzymane wyniki klasyfikacji badanego sygnału fonicznego na podstawie oceny jego cech tj. współczynników MFCC uzasadniają zastosowanie metody parametryzacji sygnału fonicznego w modelowaniu zjawisk wibroakustycznych maszyn/urządzeń, która pozwala uwzględnić efekty psychoakustyczne hałasu. Przedstawiony sposób może mieć bezpośrednie zastosowanie w zadaniach m.in. oceny źródeł dźwięku, kształtowania zdrowia pracowników, czy też opracowywania rozwiązań projektowo-konstrukcyjnych ochrony przed hałasem. Zaproponowana metoda parametryzacji sygnału fonicznego rozszerza i uzupełnia stosowane energetyczne podejście do modelowania zjawisk wibroakustycznych maszyn/urządzeń. Możliwości analizy cepstralnej opartej na charakterystyce ludzkiego słyszenia (skali Melowej) stanowiły dla autora ważny argument dla poszukiwania nowego sposobu badań modelowania zjawisk wibroakustycznych. Przedstawiony sposób pozwala na dokonanie klasyfikacji dowolnego sygnału fonicznego pod względem oceny hałasu w oparciu o Melową skalę słyszenia, dla zidentyfikowanego rodzaju źródła energii wibroakustycznej. Ważną zaletą stosowania metody parametryzacji sygnału fonicznego jest możliwość modelowania cepstrum, co pozwala na uwzględnienie percepcji hałasu i interpretację zawartych informacji w widmie sygnału fonicznego. Parametry wyznaczane w skali perceptualnej odzwierciedlają naturalne wrażenia słuchowe, co ma znaczenie dla psychoakustycznej oceny zjawisk wibroakustycznych. Opracowany sposób modelowania zjawisk wibroakustycznych opiera się na zastosowaniu oceny sygnału akustycznego/fonicznego przy wykorzystaniu metod tj.:

- metody skalowania sygnału fonicznego (subiektywna ocena wrażeń akustycznych w ramach eksperymentu psychoakustycznego),
- metody parametryzacji sygnału fonicznego (ocena cech sygnału w skali Melowej) stanowiącej obiektywny opis percypowanego hałasu.

Zakłada się, że podjęte w tym zakresie badania kontynuowane będą w obszarze stosowalności metod i modeli opisujących efekty psychoakustyczne hałasu w dziedzinie czasowo-częstotliwościowej oraz innej skali perceptualnej, np. skali Barkowej. Zaproponowany przez autora sposób modelowania zjawisk wibroakustycznych daje nowe możliwości w zakresie oceny maszyn/urządzeń na etapie:

- projektowania i konstruowania, jako dodatkowe kryterium w procesie optymalizacji cichobieżności pracy elementów,
- eksploatacji, w ramach badań diagnostycznych dla potrzeb weryfikacji stopnia zużycia poszczególnych elementów i podzespołów, a w efekcie oceny poziomu ich niezawodności.

Przetwarzane informacje w postaci pozyskanych cech z sygnału fonicznego stanowią wartość dodaną dla sygnału akustycznego i mogą być skutecznie wykorzystywane we wspomaganiu podejmowania decyzji ww. zadań.

Acknowledgement: Artykuł jest wynikiem badań prowadzonych w Katedrze Inżynierii Produkcji Wydziału Organizacji i Zarządzania Politechniki Śląskiej, które sfinansowane zostały ze środków Rektorskiego Grantu Habilitacyjnego o numerze 13/030/RGH19/0057.

Bibliografia

1. Batko W, Dąbrowski Z, Engel Z, Kiciński J, Weyna S. Nowoczesne badania procesów wibroakustycznych. Radom: Wydawnictwo Instytutu Technologii Eksploatacji –PIB, 2005.
2. Cempel Cz. Wibroakustyka stosowana. Warszawa: Państwowe Wydawnictwo Naukowe, 1989.
3. Danielewski K, Kosicka E, Mazurkiewicz D. Analysis of the correctness of determination of the effectiveness of maintenance service actions. Management and production engineering review 2018; 9(2): 20-25.
4. Directive 2002/49/Ec of the European Parliament and of the Council of 25 June 2002 relating to the assessment and management of environmental Noise. Official Journal of the European Communities – 18.07.2002.
5. Engel Z, Piechowicz J, Pleban D, Strychniewicz L. Minimalizacja przemysłowych zagrożeń wibroakustycznych. Poradnik. Warszawa: Centralny Instytut Ochrony Pracy - Państwowy Instytut Badawczy, 2005.
6. Engel Z, Zawieska Wiktor M. Hałas i drgania w procesach pracy. Źródła, ocena, zagrożenia. Warszawa: Centralny Instytut Ochrony Pracy - Państwowy Instytut Badawczy, 2010.
7. Gauchev T, Fakotakis N, Kokkinakis G. Comparative evaluation of various mfcc implementations on the speaker verification tasks. Proc. of the SPECOM 2005: 191-194.
8. Genuit K. Fiebig A. Psychoacoustics and its benefit for the soundscape approach. Acta Acustica 2006; 92: 1-7.
9. Gola A. Reliability analysis of reconfigurable manufacturing system structures using computer simulation methods. Eksploatacja i Niezawodność - Maintenance and Reliability 2019; 21(1): 90-102.
10. Jasiulewicz-Kaczmarek M., Zywica P.: The concept of maintenance sustainability performance assessment by integrating balanced scorecard with non-additive fuzzy integral. Eksploatacja i Niezawodność - Maintenance and Reliability. 2018; 20(4): 650-661.
11. Kosicka E, Kozłowski E, Mazurkiewicz D. The use of stationary tests for analysis of monitored residual processes. Eksploatacja i Niezawodność - Maintenance and Reliability 2015; 17(4): 604-609.
12. Licklider J.C.R. Correlates of the auditory stimulus. [w:] Handbook of Experimental Psychology (red. Stevens S.S.). New York: Wiley, 1966.
13. Loska A, Moczulski W, Wyczółkowski R, Dąbrowski A. Integrated system of control and management of exploitation of water supply system. Diagnostyka 2016; 17(1): 65-74.
14. Loska A. Methodology of variant assessment of exploitation policy using numerical taxonomy tools. Management Systems in Production Engineering 2015; 2(18): 98-104.
15. Makarewicz R. Dźwięk w środowisku. Poznań: Ośrodek Wydawnictw Naukowych, 1994.
16. Ozimek E. Kocinski J. Kutzner D. et al. Speech intelligibility for different spatial configurations target speech and competing noise source in a horizontal and median plane. Speech communication. 2013; (10): 1021-1032.
17. Paszkowski W. The assessment of acoustic effects of exploited road vehicles with the use of subjective features of sound. Eksploatacja i Niezawodność – Maintenance and Reliability 2019; 21(3): 522–529.

18. Paszkowski W, Kotus J, Poremski T, Kostek B. Evaluation of sound quality features on environmental noise effects - a case study applied to road traffic noise. Metrology and Measurement Systems 2018; 25(3): 517-531.
19. Paszkowski W, Sobiech M. The modeling of the acoustic condition of urban environment using noise annoyance assessment. Environmental Modeling &Assessment 2019; 24: 319-330.
20. Raimbault M, Lavandier C, Berengier M. Ambient sound assessments in or urban environments: field studies in two French cities. Applied Acoustics 2003; 64: 1341–1256.
21. Rozporządzenie Ministra Środowiska z dnia 1 października 2012 r. zmieniające rozporządzenie w sprawie dopuszczalnych poziomów hałasu w środowisku, (Dz.U. z 8 października 2012, poz. 1109).
22. Rozporządzenie Ministra Środowiska z dnia 30 października 2014r. w sprawie wymagań w zakresie prowadzenia pomiarów wielkości emisji oraz pomiarów ilości pobieranej wody,Dz.U.2014, poz. 1542.
23. Rudno-Rudzińska B. Modelowanie emisji i propagacji dźwięku do prognozowania klimatu akustycznego środowiska zurbanizowanego. Prace naukowe Instytutu Telekomunikacji i Akustyki Nr 75. Monografie nr 39. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, 1994.
24. Schomer PD. Criteria for assessment of noise annoyance. Noise Control Engineering Journal 2005; 53(4): 132-144.
25. Zieliński T. Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań. Warszawa: Wydawnictwa Komunikacji i Łączności WKŁ, 2007.
26. Zieliński T.P, Korohoda P, Rumian R. Cyfrowe przetwarzanie sygnałów w telekomunikacji. Warszawa: Wydawnictwo Naukowe PWN SA, 2014.

Dr inż. Bartosz Ceran

Wydział Inżynierii Środowiska i Energetyki
Politechnika Poznańska
ul. Piotrowo 3a, 60-965 Poznań, Polska
e-mail: bartosz.ceran@put.poznan.pl

Mgr inż. Agata Orłowska

Wydział Inżynierii Środowiska i Energetyki
Politechnika Poznańska
ul. Piotrowo 3a, 60-965 Poznań, Polska
e-mail: agata.orlowska@put.poznan.pl

Mgr inż. Krystian Krochmalny

Wydział Mechaniczno-Energetyczny
Politechnika Wrocławska
ul. C.K. Norwida 1/3, 50-370 Wrocław, Polska
e-mail: krystian.krochmalny@pwr.edu.pl

Metoda wyznaczania szybkości spadku wydajności stosu ogniw paliwowych typu PEMFC na podstawie przesunięcia charakterystyki napięciowo-prądowej

The method of determining PEMFC fuel cell stack performance decrease rate based on the voltage-current characteristic shift

Słowa kluczowe: *eksploatacja systemów ogniw paliwowych, wskaźniki eksploatacyjne stosu ogniw paliwowych, spadek wydajności stosu po latach eksploatacji, charakterystyka eksploatacyjna stosu.*

Streszczenie: Artykuł przedstawia model matematyczny przeznaczony do wyznaczenia szybkości spadku wydajności stosu ogniw paliwowych. Szybkość spadku wydajności stosu ogniw jest wyznaczana na podstawie wartości napięcia średniego stosu. Zaproponowany model wykorzystano do wyznaczenia krzywej mocy i wskaźników eksploatacyjnych stosu ogniw paliwowych o mocy nominalnej 50 kW po 14 000 h ciągłej pracy. Model wykorzystano także do wyznaczenia szybkości zmiany wartości napięcia średniego jedenastoletniego stosu ogniw paliwowych o mocy 1,2 kW. W obu badaniach wyznaczono wartości wskaźników eksploatacyjnych oraz ich różnice względem wielkości nominalnych.

Keywords: *fuel cells system exploitation, fuel cell stack exploitation indicators, stack performance decrease after years of operation, stack exploitation characteristics.*

Abstract: The article presents mathematical model designed to estimate the rate of performance decrease in fuel cell stack. The fuel cell stack performance decrease rate is determined on the basis of stack average voltage measurements. The proposed model is used to determine power curve as well as exploitation indicators of fuel cell stack with a nominal power of 50 kW after 14 000 hours of continuous operation. The model is also used to determine the average voltage drop for the eleven-year fuel cell stack with a nominal power of 1,2 kW. In both studies, the values of exploitation indicators as well as their differences in relation to nominal values are determined.

1. Wprowadzenie

Wodorowe ogniwa paliwowe z jonowymienną membraną polimerową (PEMFC- ang. Proton Exchange Membrane Fuel Cell) są uznawane za jedną z najbardziej obiecujących i perspektywicznych technologii wytwarzania energii elektrycznej i ciepła. Przewiduje się ich zastosowanie dla elektrowni dużych mocy, małych źródeł rozproszonych [14] oraz sektora transportu [12,30]. Ogniwa paliwowe mogą być eksploatowane w szerokim zakresie zmienności obciążeń elektrycznych cechując się przy tym korzystnymi wartościami wskaźników eksploatacyjnych takich jak: sprawność przetwarzania energii chemicznej paliwa na elektryczną, jednostkowe zużycie energii chemicznej paliwa, jednostkowe zużycie paliwa. Mimo to wskaźniki niezawodnościowe pracy ogniw wciąż osiągają niezadowalające wartości i są główną przyczyną zahamowania komercjalizacji tej technologii na szeroką skalę [21].

Poprawa wskaźników niezawodnościowych i określenie wskaźników eksploatacyjnych jest obecnie aktualnym i ważnym problemem. Rozwój technologii ogniw paliwowych będzie głównym motorem napędowym rozwoju sektora energetyki wodorowej oraz transportu opartego na paliwie wodorowym. Prowadzone są badania nad wykorzystaniem wodoru jako paliwa dodatkowego do silników benzynowych i wysokoprężnych [18,20]. Analiza wyników badań przeprowadzona przez autorów pracy [20] wykazała brak zasadności stosowania wodoru jako paliwa dodatkowego w silnikach o zapłonie samoczynnym. Do zasilania wodorem przystosowuje się głównie silniki o zapłonie iskrowym. W pracy [30] stwierdzono, że jest to jednak rozwiązanie tymczasowe mające na celu wstępne przygotowanie i wdrożenie infrastruktury przechowywania i rozprowadzania wodoru przed wprowadzeniem perspektywicznych ogniw paliwowych.

Ze względu na ciągle wysokie koszty ogniw paliwowych wiele badań nad tą technologią jest wykonywanych za pomocą autorskich modeli matematycznych [2, 34]. Prace dotyczące modelowania stosu ogniw paliwowych można podzielić na dwie grupy. Do pierwszej należą prace przedstawiające modele przeznaczone do optymalizacji parametrów stosu według funkcji celu takich jak: minimalizacja kosztów budowy, maksymalizacja gęstości prądu [1,4,6,15,27,32].

W artykule [4] autorzy przedstawiają trójwymiarowy wielofazowy model ogniwa paliwowego typu PEM przeznaczony do zbadania wpływu ciśnienia montażowego na rezystancję styku między warstwą dyfuzji gazu (GDL) a płytką bipolarną. Właściwy dobór wartości ciśnienia montażowego pozwala na wykonanie stosu o dłuższym czasie życia.

W pracy [6] autorzy prezentują model przeznaczony do optymalizacji wymiarów stosu ogniw paliwowych według ulepszonej wersji algorytmu optymalizacji mewy. Model pozwala na wyznaczenie charakterystyki eksploatacyjnej stosu, jednak nie uwzględnia jej przesunięcia po czasie eksploatacji. Z kolei w pracy [1] autorzy przedstawiają model stosu ogniw paliwowych o mocy 500 watów z jonowymienną membraną polimerową zaimplementowany w środowisku Matlab/Simulink. Model służy do wyznaczania wartości referencyjnej prądu elektrycznego dla dowolnego stanu ustalanego.

W pracy [27] autorzy przedstawiają procesy zachodzące w stosie ogniw paliwowych oraz opracowane modele numeryczne mające na celu minimalizację kosztów produkcji stosu i maksymalizację gęstości prądu. Modele te są przeznaczone do wspomagania projektowania stosów ogniw paliwowych.

Opracowywane są nowe techniki modelowania ogniw paliwowych np. z wykorzystaniem tzw. wykresów wiązań [32]. Prowadzone są także badania modelowe nad zwiększeniem efektywności energetycznej poprzez połączenie stosu ogniw z generatorem termoelektrycznym w układ hybrydowy [15].

Wszystkie wyżej wymienione modele matematyczne są budowane w celu wspomagania projektowania stosu ogniw paliwowych. Modele te nie uwzględniają natomiast spadku wydajności stosu po latach eksploatacji.

Do drugiej grupy publikacji dotyczących modelowania stosów należą prace, w których modeluje się procesy degradacyjne zachodzące w stosie podczas jego pracy. Autorzy skupiają się na wyznaczaniu wpływu jednego z elementów konstrukcyjnych, takich jak warstwa dyfuzyjna gazu (z ang. gas diffusion layer – GDL) [23], płytki dwubiegunowe, membrana polimerowa [26] i warstwa katalizatora na elektrodach stosu [31], na charakterystykę napięciowo-prądową. Modele te są wykorzystywane do minimalizacji źródła degradacji i zwiększenia żywotności stosu.

W pracy [26] przedstawiono numeryczny model ogniw paliwowych przeznaczony do badania mechanicznej degradacji membrany spowodowanej lokalnymi nieprężeniami. W celu zmniejszenia naprężeń autorzy sugerują zastosowanie dodatkowej uszczelki w łączeniu elektroda-elektrolit z GDL.

Kolejne prace dotyczą określenia prawdopodobieństwa degradacji danego elementu [25,35]. Autorzy artykułu [25] zaproponowali wykorzystanie drzewa błędów (z ang. faul three analysis) do wyznaczenia prawdopodobieństwa wystąpienia degradacji elementów stosu ogniw paliwowych. Z kolei w artykule [35] zaproponowano nowatorski model prognozowania degradacji ogniwa paliwowego z membraną wymiany protonów z wykorzystaniem metody filtru cząsteczkowego i ekstrapolacji danych.

Niestety wyżej wymienione modele i proponowane metody nie są nastawione na praktyczne zastosowanie, ponieważ brakuje w nich wykazania, jak postępujące procesy degradacyjne wpłyną na wydajność stosu ogniw paliwowych po wielu godzinach bądź latach eksploatacji a przede wszystkim jak zmienią się wartości wskaźników eksploatacyjnych. Znajomość zmian wartości wskaźników eksploatacyjnych będzie miała decydujący wpływ na poprawne określenie kosztów eksploatacji stosu.

Publikacje dotyczące modelowania hybrydowych systemów wytwórczych z ogniwem paliwowym [16,19,28] potwierdzają, że zapotrzebowanie na model wykazujący wpływ starzenia się stosu ogniw paliwowych na wskaźniki eksploatacyjne jest obecnie bardzo wysokie. Techniki modelowania systemów hybrydowych przedstawione w wyżej wymienionych pracach nie uwzględniają spadku wydajności stosu po latach eksploatacji. W konsekwencji, w takim systemie hybrydowym może dojść do niezbilansowania energii [24], co będzie skutkowało zwiększeniem poboru z sieci elektroenergetycznej [9].

Analogicznie, jeśli ogniwo paliwowe jest wykorzystywane w sektorze transportu [3,11,29], nieuwzględnienie spadku wydajności stosu może doprowadzić do błędnego oszacowania zasięgu pojazdu lub zapotrzebowania na paliwo, co spowoduje błędne oszacowanie kosztów eksploatacji pojazdu.

W celu poprawnego określenia kosztów eksploatacyjnych pojazdu lub systemu hybrydowego wytwarzania energii elektrycznej z ogniwem paliwowym należy budować modele matematyczne, które umożliwią określenie spadku wydajności ogniw po latach eksploatacji.

Głównym celem niniejszego artykułu jest prezentacja modelu matematycznego pozwalającego na wyznaczenie szybkości spadku wydajności stosu ogniw paliwowych oraz przesunięcia charakterystyki eksploatacyjnej stosu względem charakterystyki dla danych katalogowych po zadanym czasie użytkowania. Model pozwala na wyznaczenie wartości wskaźników eksploatacyjnych tj. sprawności elektrycznej, jednostkowego zużycia energii chemicznej paliwa na produkcję energii elektrycznej, jednostkowego zużycia paliwa na produkcję energii elektrycznej i porównanie ich z wartościami wskaźników odpowiadającymi charakterystyce katalogowej stosu.

2. Model matematyczny ogniwa paliwowego

Opisywane w literaturze modele matematyczne ogniw paliwowych pozwalają na wyznaczenie charakterystyki eksploatacyjnej dla warunków znamionowych (parametrów katalogowych). Proponowane modele nie uwzględniają spadku wydajności stosu po latach życia i eksploatacji.

Zaproponowana przez autorów metoda analizy pozwala na wyznaczanie charakterystyki eksploatacyjnej stosu ogniw paliwowych dla danych znamionowych, oraz charakterystyki po „n” latach eksploatacji. Z drugiej strony proponowany algorytm można wykorzystać do określenia szybkości spadku wydajności stosu na podstawie przesunięcia między charakterystyką katalogową a charakterystyką stosu po zadanym okresie użytkowania.

Metoda pozwala na szybkie wyznaczanie sprawności oraz pozostałych wskaźników eksploatacyjnych stosu ogniw paliwowych.

Sprawność przetwarzania energii chemicznej paliwa na elektryczną przez stos ogniw paliwowych można przedstawić za pomocą wzoru (1) [7]:

$$\eta_{el} = \frac{P_{el}}{\dot{n}_{H_2} \cdot Q_{W_{H_2}}} = \frac{U \cdot I}{\dot{n}_{H_2} \cdot Q_{W_{H_2}}} \tag{1}$$

gdzie: P_{el} – moc elektryczna stosu ogniw paliwowych, U – napięcie stosu ogniw paliwowych, I – natężenie prądu, $\dot{n}_{H_2}$- molowy strumień wodoru, $Q_{W_{H2}}$ – wartość opałowa wodoru odniesiona odpowiednio do 1 mola wodoru.

Wskaźnik jednostkowego zużycia energii chemicznej paliwa na produkcję energii elektrycznej przez stos definiuje wzór (2):

$$q_{el} = \frac{\dot{n}_{H_2} \cdot Q_{W_{H_2}}}{P_{el}} [\frac{kJ}{kJ}] \tag{2}$$

Wskaźnik jednostkowego zużycia paliwa na produkcję energii elektrycznej można obliczyć za pomocą wzoru (3):

$$b_{el} = \frac{V_{H_2}}{P_{el}} [\frac{Nm^3 H_2}{kWh}] \tag{3}$$

gdzie: V_{H_2} – objętościowy strumień wodoru.

Należy zaznaczyć, że producenci systemów ogniw paliwowych typu PEMFC w specyfikacji technicznej podają dwie wartości sprawności przetwarzania energii chemicznej paliwa na elektryczną: wartość sprawności odniesioną do wartości opałowej wodoru (ang. LHV – Low Heating Value) oraz wartość sprawności odniesioną do ciepła spalania wodoru (ang. HHV – High Heating Value). Analiza przedstawiona w tej pracy jest prowadzona w odniesieniu do wartości opałowej wodoru, ponieważ wartość wskaźnika q_{el} definiuje się, w klasycznej teorii eksploatacji źródeł wytwórczych, w odniesieniu do tej wielkości.

Definiowanie sprawności ogniwa paliwowego w odniesieniu do wartości opałowej pozwala na porównywanie jej wartości z innymi technologiami wytwarzania energii elektrycznej wykorzystującymi energię chemiczną paliwa, np. konwencjonalne bloki parowe, bloki gazowe i gazowo-parowe, silniki gazowe czy technologie wykorzystujące biomasę.

Wzór na strumień molowy wodoru można wyznaczyć na podstawie II prawa elektrolizy Faradaya (4):

$$q = n \cdot z \cdot F \tag{4}$$

gdzie: q [C] – ładunek elektryczny, n – liczba moli gazu, z [-] liczba elektronów koniecznych do uwolnienia cząsteczki dla H_2 = 2, tzn. 2 mole elektronów są potrzebne do uwolnienia 1 mola H_2, dla O_2 = 4, F [C/mol] – stała Faradaya.

Dzieląc równanie (4) przez czas t otrzymuje się zależność na natężenie prądu elektrycznego (5):

$$I = \frac{n}{t} \cdot z \cdot F \tag{5}$$

Z równania (5) można wyznaczyć strumień molowy oznaczony jako $\dot{n}$ (6):

$$\dot{n} = \frac{n}{t} = \frac{I}{z \cdot F} \tag{6}$$

Dla stosu ogniw paliwowych strumień molowy wodoru jest proporcjonalny do liczby celek wchodzących w skład stosu, stąd:

$$\dot{n}_{H_2} = \frac{I \cdot n_{celek}}{z \cdot F} \tag{7}$$

gdzie: n_{celek} – liczba celek (pojedynczych ogniw), z których zbudowany jest stos

W teorii ogniw paliwowych praktyczne zastosowanie ma wielkość nazwana potencjałem termoneurtralnym, który jest zdefiniowany wg wzoru (8) [22]:

$$E_t^0 = -\frac{\Delta H^0_{H_20(g)}}{z \cdot F} \tag{8}$$

gdzie: E_t^0 – potencjał termoneutralny [V], $\Delta H^0_{H_20(g)}$ – standardowa entalpia tworzenia wody w fazie gazowej [kJ/mol], indeks “0” oznacza warunki standardowe (T = 298 K, p = 10^5 Pa).

Potencjał termoneutralny jest teoretyczną wartością napięcia jakie osiągnie ogniwo paliwowe przy teoretycznym założeniu, że 100 % strumienia dostarczonej energii chemicznej zostanie przekształcone w energię elektryczną.

Standardowa entalpia tworzenia wody w fazie gazowej $\Delta H_{H2O(g)}$ energetycznie odpowiada wartości opałowej wodoru, przy założeniu, że woda jest produktem w stanie gazowym (9) [17]:

$$-\Delta H^0_{H_20(g)} = Q_{W_{H2}} \tag{9}$$

Podstawiając równania (4), (8), (9) do równania (1) wzór na sprawność wytwarzania energii elektrycznej przez stos ogniw paliwowych przyjmuje postać wzoru (10):

$$\eta_{el} = \frac{U}{n_{celek} \cdot E_t^0} \tag{10}$$

Wartość średniego napięcia stosu ogniw paliwowych zdefiniowana jest jako stosunek napięcia stosu do liczby celek (pojedynczych ogniw) tworzących stos:

$$U_{av} = \frac{U}{n_{celek}} \tag{11}$$

Po podstawieniu zależności (11) do wzoru (10) wzór na sprawność stosu można przedstawić z pomocą zależności:

$$\eta_{el} = \frac{U_{av}}{E_t^0} \tag{12}$$

Wzór (12) na sprawność przetwarzania energii chemicznej paliwa na elektryczną przez stos jest bardzo praktyczny, ponieważ do wyznaczenia sprawności wystarczy pomiar napięcia stosu i znajomość liczby pojedynczych ogniw (celek) tworzących stos. Nie potrzeba natomiast dokonywać pomiarów strumienia zużywanego wodoru.

Miarą zmiany sprawności stosu jest zmiana wartości średniego napięcia stosu. Zmiana sprawności jest zgodnie z zależnością (13) wprost proporcjonalna do zmiany wartości średniego napięcia stosu:

$$\Delta\eta_{el} = \frac{\Delta U_{av}}{E_t^0} \tag{13}$$

Zmianę wartości średniego napięcia stosu definiuje wzór (14):

$$\Delta U_{av} = U_{av} - \frac{dU_{av}}{dt} \cdot t \tag{14}$$

gdzie: $\frac{dU_{av}}{dt}$ – szybkość zmiany wartości średniego napięcia stosu w czasie [μV/h], t – czas życia stosu ogniw paliwowych [h].

Zmianę mocy stosu można wyznaczyć na postawie zależności (15):

$$\Delta P_{el} = \Delta\eta_{el} \cdot \dot{n}_{H_2} \cdot Q_{WH_2} \tag{15}$$

Przedstawiony powyżej algorytm pozwala w szybki sposób na określenie szybkości zmiany wartości średniego napięcia stosu, a co za tym idzie szybkości spadku generowanej mocy elektrycznej i sprawności przetwarzania energii chemicznej paliwa na elektryczną a także wzrostu wartości wskaźników jednostkowego zużycia paliwa i jednostkowego zużycia energii chemicznej przez stos.

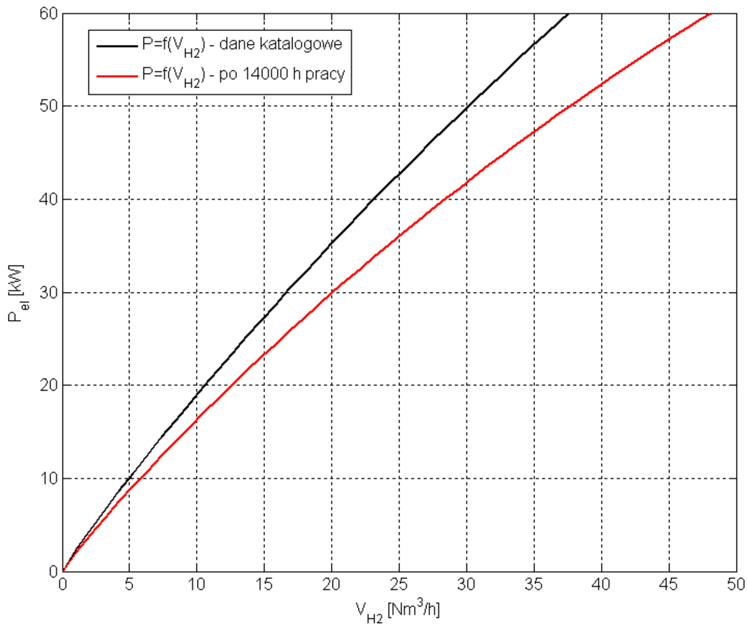
Dane wejściowe modelu stanowią parametry katalogowe stosu oraz wyznaczona na podstawie pomiarów wartość napięcia stosu. Na podstawie wzorów (1) – (15) wyznaczane są wartości wskaźników eksploatacyjnych stosu oraz wykreślane charakterystyki eksploatacyjne dla parametrów katalogowych oraz po zadanym okresie użytkowania. W celu wyznaczenia szybkości spadku wydajności stosu algorytm wykonuje szereg symulacji dla różnych wartości szybkości zmian średniego napięcia stosu z zadanym krokiem i ocenia, z zadaną dokładnością, dopasowanie charakterystyki rzeczywistej do symulowanej.

3. Wyznaczanie wskaźników eksploatacyjnych na podstawie szybkości zmiany wartości średniego napięcia stosu

Zmianę wartości średniego napięcia stosu ΔU_{av} można aproksymować funkcją liniową [33]. Szybkość zmiany wartości średniego napięcia w czasie dla systemów o mocy rzędu kilkudziesięciu kilowatów zawiera się, na podstawie doświadczeń eksploatacyjnych, w zakresie od około 3 do 9 μV/h [33]. Wartość ta zależy od wielu czynników, takich jak kultura eksploatacji stosu (przestrzeganie procedur ruchowych, rozruchu, odstawiania stosu), warunki pracy stosu (warunki atmosferyczne, temperatura otoczenia), charakter pracy stosu (ciągła, przerywana) itp.

Pilotażowa instalacja zbudowana z ogniw PEM o mocy 50 kW (12 stosów o mocy 4,2 kW) w Niderlandach, w miejscowości Delfzijl przepracowała bez przerw 14 000 h. Pomiary w czasie eksploatacji wykazały, że wartość średniego napięcia stosu zmniejszała się z szybkością średnią 8 μV/h [33].

Bazując na równaniach (1) – (15) oraz autorskim kodzie opracowanym w środowisku Matlab przeprowadzono symulacje w celu wyznaczenia krzywej mocy stosu oraz wartości wskaźników eksploatacyjnych. Wyniki symulacji przedstawiono na rysunku 1 oraz w tabeli 1.



Rys. 1. Wpływ zmiany wartości U_{av} na krzywą mocy stosu ogniw paliwowych o mocy 50 kW – badania symulacyjne

Tabela 1. Wskaźniki eksploatacyjne stosu ogniw paliwowych 50 kW – badania symulacyjne

	Dane znamionowe	**Po 14 000 h eksploatacji**
P [kW]	50	42,4
ΔP [kW]	-7,6	
η_{el} [-]	0,55	0,46
$\Delta\eta_{el}$ [-]	-0,09	
q_{el} [kJ/kJ]	1,81	2,16
Δq_{el} [kJ/kJ]	0,35	
b_{el} [Nm^3H_2/kWh]	0,60	0,72
Δb_{el} [Nm^3H_2/kWh]	0,12	

Po 14 000 h ciągłej pracy sprawność wytwarzania energii elektrycznej przez instalację zmniejszyła swoją wartość o 16,36% spadając z 55% do wartości 46%. Efektem spadku sprawności jest spadek wartości mocy elektrycznej generowanej przez stos oraz zwiększenie wartości wskaźników jednostkowego zużycia energii chemicznej paliwa oraz jednostkowego zużycia paliwa. Moc nominalna stosu zmniejszyła się o 15,2%. Oznacza to, że dodatkowe 7,6 kW strumienia energii chemicznej wodoru zostanie zamienione przez instalację na moc cieplną kosztem wartości generowanej mocy elektrycznej. Wartość wskaźnika q_{el} wzrosła o 19,3% w stosunku do wartości nominalnej, natomiast wartość wskaźnika b_{el} wzrosła o 20%.

4. Wyznaczanie wartości szybkości zmiany napięcia średniego stosu na podstawie pomiaru charakterystyki eksploatacyjnej

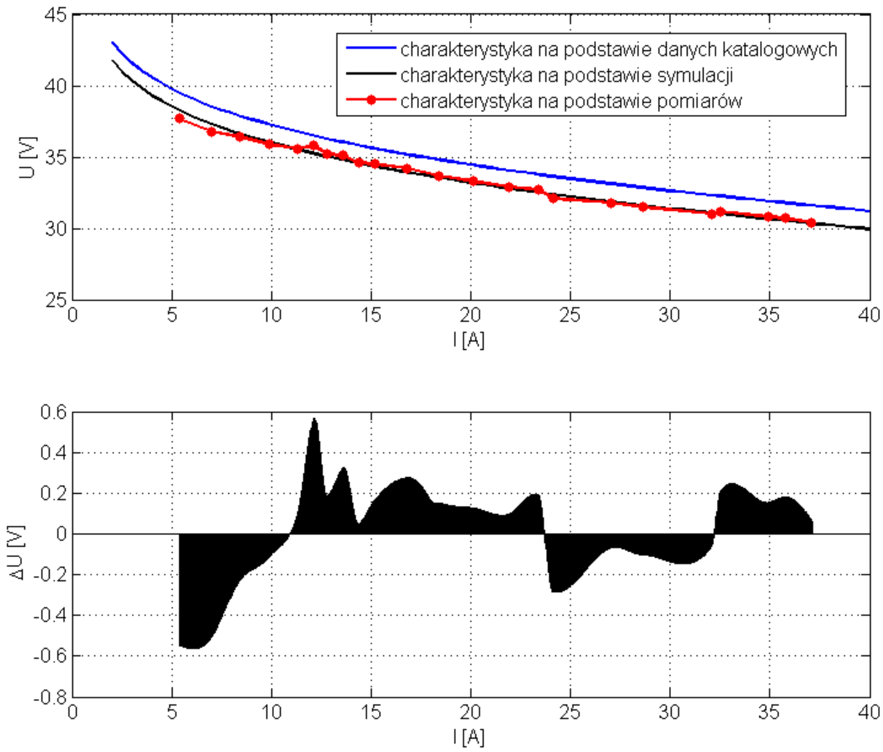
W celu weryfikacji modelu wykorzystano go do dopasowania charakterystyki napięciowo-prądowej stosu ogniw paliwowych o mocy 1,2 kW i wyznaczenia szybkości zmiany wartości napięcia średniego stosu. Badany stos ogniw paliwowych jest elementem systemu treningowego NEXA, który znajduje się w laboratorium konwersji energii Politechniki Wrocławskiej.

System NEXA jest urządzeniem przeznaczonym do awaryjnego zasilania urządzeń stało i przemiennoprądowych. Oprócz stosu o mocy 1,2 kW system zbudowany jest z [8]:

- układu zasilania wodorowego – butle ze sprężonym wodorem 20 MPa, reduktor ciśnienia wodoru, regulator ciśnienia wodoru, zawór nadmiarowego ciśnienia, zawór elektromagnetyczny odcinający dopływ paliwa podczas wyłączania systemu, detektor wycieku wodoru,
- układu zasilania powietrznego – dmuchawa typu Roots,
- układu chłodzenia stosu – stos ogniw w systemie NEXA jest chłodzony powietrzem za pomocą wentylatora chłodzącego,
- elektronicznego system sterowania – komputer sterujący, czujniki pomiarowe.

Wiek stosu ogniw paliwowych wynosi 11 lat. Na podstawie przeprowadzonej symulacji określona wartość szybkości zmiany napięcia średniego wynosi 0,34 μV/h. Dla tej wartości uzyskuje się najlepsze dopasowanie charakterystyki symulowanej do mierzonej. Niska, w porównaniu z systemem 50 kW, wartość szybkości zmiany napięcia średniego jest spowodowana tym, że stos ogniw 1,2 kW składa się z 46 celek. Dodatkowo system NEXA pracuje w trybie przerywanym, jest wykorzystywany tylko do celów badawczych i dydaktycznych.

Na rysunku 2a przedstawiono charakterystyki stosu ogniw paliwowych: nominalną, charakterystykę wyznaczoną na podstawie pomiarów (pomiar wykonany w marcu 2020) oraz charakterystykę uzyskaną na podstawie symulacji. Rysunek 2b przedstawia dopasowanie charakterystyki mierzonej do symulowanej.



Rys. 2. a) Charakterystyki stosu ogniw paliwowych: katalogowa, symulowana, zmierzona, b) ΔU między charakterystyką symulowaną a zmierzoną.

Odchylenia widoczne na charakterystyce pomiarowej są efektem działania układu płukania anody w czasie wykonywania pomiarów [5]. W tym punkcie różnica między wielkością mierzoną a symulowaną wynosi około 0,6 V. W większości przedziału obszaru omowego [10,13] charakterystyki napięciowo – prądowej różnica między punktami wynosi około 0,2 V.

W tabeli 2 przedstawiono porównanie wskaźników eksploatacyjnych obliczonych na podstawie danych katalogowych oraz na podstawie wyników symulacji.

Tabela 2. Wskaźniki eksploatacyjne stosu ogniw paliwowych 1,2 kW

	Dane znamionowe	**Po 11 latach**
P_{el} [kW]	1,2	1,16
ΔP [kW]	-0,040	
η_{el} [-]	0,550	0,521
$\Delta\eta_{el}$ [-]	-0,029	
q_{el} [kJ/kJ]	1,840	1,920
Δq_{el} [kJ/kJ]	0,080	
b_{el} [Nm^3H_2/kWh]	0,610	0,640
Δb_{el} [Nm^3H_2/kWh]	0,030	

Niska wartość szybkości zmiany napięcia średniego stosu skutkuje mniejszymi zmianami wartości mocy stosu i wskaźników eksploatacyjnych. Moc nominalna stosu ogniw paliwowych po 11 latach spadła z 1,2 kW do 1,16 kW, zmniejszając swoją wartość o 3,33%. Sprawność przemiany energii chemicznej wodoru na energię elektryczną zmalała o 5,27%. Wartość wskaźnika jednostkowego zużycia energii chemicznej paliwa na produkcję energii elektrycznej wzrosła o 4,35%. Jednostkowe zużycie paliwa na produkcję jednej kWh energii elektrycznej wzrosło o 4,92%.

Mając na uwadze fakt, że stos pracuje w warunkach laboratoryjnych z zachowaniem wysokiej kultury eksploatacji wyniki należy uznać za poprawne.

5. Podsumowanie

Aktualizacja charakterystyk eksploatacyjnych ogniw paliwowych z jonowymienną membraną polimerową jest kluczowym zagadnieniem w momencie rozwoju technologii opartych na wodorze tj. rozproszonej generacji energii elektrycznej oraz elektromobilności.

Zaproponowana metoda pozwala w łatwy i szybki sposób określić spadek wydajności stosu ogniw paliwowych, co pozwala na wyznaczenie aktualnych wartości wskaźników eksploatacyjnych stosu. Z kolei, wyznaczone wartości wskaźników eksploatacyjnych po latach użytkowania stosu pozwalają na bardziej dokładne oszacowanie kosztów eksploatacji systemu ogniw paliwowych.

Proponowana metoda stanowi przydatne narzędzie do przeprowadzania analizy studium wykonalności danego projektu (np. hybrydowy system wytwórczy z magazynem energii w postaci wodoru, samochód lub autobus na wodór), na podstawie której, inwestor będzie w stanie dokładniej ocenić ryzyko związane z danym przedsięwzięciem, urzeczywistniając szacunki techniczno-finansowe. Proponowany model może stanowić ułatwienie w planowaniu długoterminowej pracy stosów ogniw paliwowych zarówno w układach generacji rozproszonej jak i w pojazdach wodorowych.

Zaletami proponowanej metody są jej prostota, krótki czas obliczeń oraz fakt, że do określenia wydajności stosu ogniw paliwowych wystarczy wykonać pomiar napięcia generowanego przez stos.

Przeprowadzone symulacje wykazały, że wartość spadku wydajności stosu zależy od jego mocy nominalnej (pośrednio od liczby celek) oraz trybu i warunków pracy (praca ciągła, praca przerywana). Przedstawiony w artykule model stanowi przyczynek do rozwoju metod aktualizacji charakterystyk eksploatacyjnych stosów ogniw paliwowych. Zagadnienia te będą miały coraz większe znaczenie ze względu na spodziewaną decentralizację struktury sektora wytwórczego systemu elektroenergetycznego, rozwój hybrydowych systemów wytwórczych ze stosami ogniw paliwowych oraz elektromobilności. Spadek wydajności stosu ogniw paliwowych w trakcie lat eksploatacji to aspekt niezwykle istotny dla inwestorów zainteresowanych nowymi technologiami wodorowymi.

References

1. Abdelnasir O, Smith D, Alaswad A, Amiri A, Sodre JR, Lucchesi A. Proton-exchange membrane (PEM) fuel cell system mathematical modelling. Paper presented at SDEWES - 14th Conference on Sustainable Development of Energy, 2019; Water and Environment Systems, Dubrovnik, Croatia.
2. Abdul Rasheed RK, Liao Q, Caizhi Z, Chan SH. A review on modelling of high temperature proton exchange membrane fuel cells (HT-PEMFCs). International Journal of Hydrogen Energy 2017; 42(5): 3142-3165, https://doi.org/10.1016/j.ijhydene.2016.10.078.
3. Álvarez Fernández R, Corbera Caraballo S, Beltrán Cilleruelo F, Lozano JA. Fuel optimization strategy for hydrogen fuel cell range extender vehicles applying genetic algorithms. Renewable and Sustainable Energy Reviews 2018; 81(1): 655-668, https://doi.org/10.1016/j.rser.2017.08.047.
4. Atyabi SA, Afshari E, Wongwises S, Yan WM, Hadjadj A, Shadloo MS. Effects of assembly pressure on PEM fuel cell performance by taking into accounts electrical

and thermal contact resistances. Energy 2019; 179: 490-501, https://doi.org/10.1016/j.energy.2019.05.031.

5. Barbir F. PEM Fuel Cell: Theory and Practice. New York: Elsevier Academic Press, 2005, 99–113.
6. Cao Y, Li Y, Zhang G, Jermsittiparsert K, Razmjooy N. Experimental modeling of PEM fuel cells using a new improved seagull optimization algorithm. Energy Reports 2019; 5: 1616-1625, https://doi.org/10.1016/j.egyr.2019.11.013.
7. Ceran B, Bernstein P.A. Operational characteristics of proton exchange membrane (PEM) fuel cells, Przegląd Elektrotechniczny 2014; 10: 102 – 105, 2014, doi: 10.12915/pe.2014.10.26.
8. Ceran B, Długosz J, Kruczek-Pawlak H. Analiza energetyczna stosu ogniw paliwowych z jonowymienna membrana polimerowa PEMFC, Poznań University of Technology Academic Journals, 2016; 86: 301-312.
9. Ceran B, Orłowska A. The Impact of Power Source Performance Decrease in a PV/WT/FC Hybrid Power Generation System on the Result of a Multi-Criteria Analysis of Load Distribution. Energies 2019; 12(18): 3453, https://doi.org/10.3390/en12183453.
10. Ceran B. Charakterystyki eksploatacyjne stosu ogniw paliwowych typu PEMFC. Polityka Energetyczna – Energy Policy Journal 2014; 17(3): 135–146.
11. Chandrasekar C, Amruth Kumar L. A Novel Approach on Range Prediction of a Hydrogen Fuel Cell Electric Truck. SAE Technical Paper 2019; 2019-28-2514.
12. Chmielniak T, Lepszy S, Mońka P. Energetyka wodorowa – podstawowe problemy. Polityka Energetyczna – Energy Policy Journal 2017; 20(3): 55–66.
13. Dudek M, Celowski P, Lis B, Raźniak A, Dudek P. Laboratoryjny generator energii elektrycznej o mocy 360 W zawierający niskotemperaturowy stos ogniw paliwowych PEMFC chłodzony za pomocą medium ciekłego. Przegląd Elektrotechniczny 2016; 10: 235 – 242, 2014, doi:10.15199/48.2016.10.54.
14. Gharehpetian GB, Mohammad Mousavi Agah S. Distributed Generation Systems. Design, Operation and Grid Integration. Butterworth-Heinemann. An imprint of Elsevier, 2017, https://doi.org/10.1016/B978-0-12-804208-3.09993-3.
15. Guo X, Zhang H, Wang J, Zhao J, Wang F, Miao H, Yuan J, Hou S. A new hybrid system composed of high-temperature proton exchange fuel cell and two-stage thermoelectric generator with Thomson effect: Energy and exergy analyses. Energy 2020; 195: 117000, https://doi.org/10.1016/j.energy.2020.117000.
16. Hosseinalizadeh R, Shakouri GH, Amalnick MS, Taghipour P. Economic sizing of a hybrid (PV–WT–FC) renewable energy system (HRES) for stand-alone usages by an optimization-simulation model: Case study of Iran. Renewable and Sustainable Energy Reviews 2016; 54: 139-150, https://doi.org/10.1016/j.rser.2015.09.046.
17. Kabza A. Fuel Cell Formulary, www.kabza.de, dostęp: 31.03.2020.
18. Keršys A, Kalisinskas D, Pukalskas S, Vilkauskas A, Keršys R, Makaras R. Investigation of the influence of hydrogen used in internal combustion engines on exhaust emission. Eksploatacja I Niezawodnosc – Maintenance and Reliability 2013; 15 (4): 384–389.
19. Khan A, Javaid N. Optimum Sizing of PV-WT-FC-DG Hybrid Energy System using Teaching Learning-Based Optimization. International Conference on Frontiers of Information Technology (FIT) Islamabad, Pakistan, 2019; 127-1275, 10.1109/FIT47737.2019.00033.
20. Kruczyński S, Ślężak M, Gis W, Orliński P. Evaluation of the impact of combustion hydrogen addition on operating properties of self-ignition engine. Eksploatacja i

Niezawodność – Maintenance and Reliability 2016; 18 (3): 343–347, http://dx.doi.org/10.17531/ein.2016.3.4.
21. Mayur M, Gerard M, Schott P, Bessler WG. Lifetime Prediction of a Polymer Electrolyte Membrane Fuel Cell under Automotive Load Cycling Using a Physically-Based Catalyst Degradation Model. Energies 2018; 11: 2054, doi:10.3390/en11082054.
22. O'Hayre R, Cha SW, Colella W, Prinz FB. Fuel Cell Fundamentals, 3rd Edition, Wiley 2016.
23. Park J, Oh H, Ha T, Lee YI, Min K. A review of the gas diffusion layer in proton exchange membrane fuel cells: Durability and degradation. Applied Energy 2015; 155: 866-880, https://doi.org/10.1016/j.apenergy.2015.06.068.
24. Paska J. Chosen aspects of electric power system reliability optimization. Eksploatacja I Niezawodnosc – Maintenance and Reliability 2013; 15 (2): 202–208.
25. Placca L, Kouta R. Fault tree analysis for PEM fuel cell degradation process modelling. International Journal of Hydrogen Energy 2011; 36(19): 12393-12405, https://doi.org/10.1016/j.ijhydene.2011.06.093.
26. Qiu D, Peng L, Liang P, Yi P, Lai X. Mechanical degradation of proton exchange membrane along the MEA frame in proton exchange membrane fuel cells. Energy 2018; 165: 210-222, https://doi.org/10.1016/j.energy.2018.09.136.
27. Secanell M, Jarauta A, Kosakian A, Sabharwal M, Zhou J. PEM Fuel Cells, Modeling. 2017; Springer, New York, NY.
28. Shayeghi H, Shahryari E, Moradzadeh M, Siano P. A Survey on Microgrid Energy Management Considering Flexible Energy Sources. Energies 2019; 12(11): 2156, https://doi.org/10.3390/en12112156.
29. Sprik S, Thornton MJ, Brooks K, Tamburello DA. Performance Modeling of Materials-Based Hydrogen Storage Systems for Automotive Applications. 2017 AIChE Annual Meeting, 29 October - 3 November 2017, Minneapolis, Minnesota.
30. Sroka ZJ. Durability of engine components due to alternative fuels. Eksploatacja I Niezawodnosc – Maintenance and Reliability 2007; 4(36): 9-15.
31. Thangavelautham J. Degradation in PEM fuel cells and mitigation strategies using system design and control. T. Taner (Ed.), Proton Exchange Membrane Fuel Cell, London: Intech Open Ltd, 2018, pp. 63-95, 10.5772/intechopen.72208.
32. Vasilyev A, Andrews J, Jackson LM, Dunnett SJ, Davies B. Component-based modelling of PEM fuel cells with bond graphs. International Journal of Hydrogen Energy 2017; 42; 29406-29421, https://doi.org/10.1016/j.ijhydene.2017.09.004.
33. Verhage A, Gerits J, Manders T. Duration Tests of PEM Fuel Cells in a 50 kW Pilot Power Plant. In Proceedings of the 18th World Hydrogen Energy Conference (WHEC), Essen, Germany, 2010; 16–20 May: 63–67.
34. Wu H-W. A review of recent development: Transport and performance modeling of PEM fuel cells. Applied Energy 2016; 165: 81-106, https://doi.org/10.1016/j.apenergy.2015.12.075.
35. Zhou D, Wu Y, Gao F, Breaz E, Ravey A, Miraoui A. Degradation Prediction of PEM Fuel Cell Stack Based on Multiphysical Aging Model With Particle Filter Approach. IEEE Transactions on Industry Applications 2017; (53)4: 4041-4052, 10.1109/TIA.2017.2680406.

dr hab. inż. Leszek UŁANOWICZ, prof. ITWL

Instytut Techniczny Wojsk Lotniczych,
ul. Księcia Bolesława 6, 01-494 Warszawa, Polska
E-mail: leszek.ulanowicz@itwl.pl

dr inż. Grzegorz JASTRZĘBSKI

Instytut Techniczny Wojsk Lotniczych,
ul. Księcia Bolesława 6, 01-494 Warszawa, Polska
E-mail: grzegorz.jastrzebski@itwl.pl

dr inż. Paweł SZCZEPANIAK

Instytut Techniczny Wojsk Lotniczych,
ul. Księcia Bolesława 6, 01-494 Warszawa, Polska
E-mail: pawel.szczepaniak@itwl.pl

Metoda do szacowania trwałości lotniczych napędów hydraulicznych

Słowa kluczowe: lotnictwo, trwałość, napęd hydrauliczny, pompa hydrauliczna, stan techniczny

Streszczenie: W dotychczasowej praktyce szacowania trwałości zespołów lotniczych napędów hydraulicznych stosowany jest wariant, który wymaga prowadzenia długotrwałych badań zespołów napędu do czasu ich przejścia w stan niezdatności. Badania tego typu, umożliwiające szacowanie trwałości a posteriori, są kosztowne i długotrwałe. Istnieje więc potrzeba poszukiwania nowych strategii szacowania trwałości. W artykule zaprezentowano metodę szacowania trwałości zespołu napędu hydraulicznego opartą o kontrolę jego zmiany stanu technicznego. Kontrola stanu technicznego umożliwia wykrycie we właściwym czasie stanu przed awaryjnego zespołu hydraulicznego. Novum metody jest wykorzystanie, do wykrycie stanu przed awaryjnego zespołu, zasady wyznaczania uprzedzających tolerancji parametru kontrolnego. Tolerancje uprzedzające stanowią zbiór wartości parametru kontrolnego zawartych między poziomami granicznym i przed awaryjnym (dopuszczalnym). Intensywność wyczerpywania się trwałości (intensywności starzenia, zużywania) ma losowy charakter. W artykule przedstawiono stochastyczny opis zmiany parametru kontrolnego oraz wynikające z niego empiryczne zależności funkcji gęstości prawdopodobieństwa czasu przeprowadzania sprawdzeń parametru kontrolnego (okresowość kontroli) i funkcji gęstości prawdopodobieństwa zmiany wartości parametru kontrolnego. Opisano wzajemne związki obu tych funkcji. Przedstawiono zależności umożliwiające wyznaczenie wartości dopuszczalnej parametru kontrolnego i okresowość sprawdzeń parametru kontrolnego po przekroczeniu wartości dopuszczalnej. Zaprezentowano przykład szacowania trwałości tłoczkowej pompy hydraulicznej z samolotu użytkowanego w Siłach Zbrojnych RP. Dla wybranych parametrów kontrolnych pompy hydraulicznej wyznaczono ich wartości dopuszczalne oraz czas pierwszej kontroli parametru kontrolnego po przekroczeniu wartości dopuszczalnej. Zaprezentowana metoda wiąże trwałość z fizycznymi mechanizmami zużywania się zespołów. Przedstawiona metoda może być wykorzystana w pracach mających na celu określanie zasobu pracy urządzeń technicznych. Umożliwia ona użytkowanie urządzeń technicznych według strategii stanu technicznego z kontrolowaniem parametrów.

1. Wstęp

Problem szacowania trwałości zespołów lotniczego napędu hydraulicznego stanowi szerokie zagadnienie prognozowania w fazie konstruowania ich eksploatacyjnych zachowań, jak również prognozowania zmiany ich stanu technicznego w fazie eksploatacji. Doświadczenia z eksploatacji lotniczych napędów hydraulicznych w statkach powietrznych (SP) wskazują, że po wykorzystaniu ustalonej przez producenta trwałości normatywnej większość zespołów hydraulicznych ma jeszcze pewien zasób pracy, który może być wykorzystany [21, 24]. Może to świadczyć o tym, że na etapie projektowania zespołów hydraulicznych niewłaściwie zidentyfikowano warunki ich pracy i narzucono nieadekwatne ograniczenia przy szacowaniu ich trwałości [14]. Istnieje więc potrzeba technicznego i naukowego poszukiwania metod szacowania trwałości korygujących przyjęte założenia projektowe z jednoczesnym zachowaniem funkcjonalności i efektów działania zespołu hydraulicznego.

Na podstawie dostępnej literatury można sobie wyrobić pewien pogląd odnośnie do ogólnych zasad szacowania trwałości zespołów hydraulicznych, przyjętych przez różne ośrodki naukowo-badawczo-produkcyjne [1, 14].

Dotychczasowa praktyka szacowania trwałości zespołów lotniczego napędu hydraulicznego jest wielowątkowa i wielokierunkowa. Główny kierunek szacowania trwałości opiera się o zasadę, że na podstawie danych z badań laboratoryjnych i stanowiskowych można oszacować trwałość zespołu w odpowiednich warunkach eksploatacyjnych [7, 23]. Drugim kierunkiem, uzupełniającym główny kierunek jest szacowanie trwałości oparte na badaniach niezawodności eksploatacyjnej zespołu [1,13]. Oba kierunki wykorzystują do projektowania zespołów koncepcje bezpiecznej trwałości.

Pierwszy kierunek szacowania trwałości wymaga prowadzenia długotrwałych i kosztownych badań zespołów hydraulicznych do czasu ich przejścia w stan niezdatności [2, 7]. W podejściu tym na etapie projektowania przeprowadza się próby zużyciowe zespołów hydraulicznych [4, 18]. Próby te wykonywane są wyłącznie w warunkach stanowiskowych [5, 11]. Mają one na celu sprawdzenie założonej odporności hydraulicznych par precyzyjnych badanego zespołu [2, 22]. Próby te prowadzi się wg specjalnie opracowanych programów dla danego zespołu, z reguły przewidujących przyspieszony ich tryb i ostrzejsze warunki obciążeń od występujących w eksploatacji [14]. Realizuje się je do czasu uszkodzenia się zespołu hydraulicznego. Na ogół program prób, ustalany w trakcie opracowywania danego zespołu hydraulicznego, przewiduje realizację szeregu jednakowych, kolejno po sobie następujących etapów, z których każdy składa się z szeregu podetapów o różnych wartościach parametrów obciążeń badanego zespołu, realizowanych w określonym czasie a więc w określonej liczbie cykli obciążeń [9, 12]. Widać stąd, że czasokres badań jest bardzo długi i stąd badania te są również kosztowne. Próby zużyciowe nie uwzględniają jednak wszystkich wymuszeń eksploatacyjnych, gdyż samo odtworzenie na stoisku rzeczywiście występujących w eksploatacji obciążeń badanego zespołu jest dużym problemem. Zasady ustalania każdorazowego programu prób są też zagadnieniem niezmiernie skomplikowanym i czasochłonnym. Rozrzut wyników eksperymentalnych badań zużyciowych przy szacowaniu trwałości zespołu hydraulicznego jest podstawą do wprowadzania współczynnika bezpieczeństwa, czyli niemianowanego stosunku wartości niebezpiecznej do wartości dopuszczalnej. Najczęściej współczynnik bezpieczeństwa uwzględnia: ewentualną nieadekwatność programu prób do rzeczywistych warunków pracy zespołu w warunkach eksploatacji [2], dostępność dla kontroli miejsca zużycia [19], charakter postępującego zniszczenia oraz prędkość destrukcji [14], stopień wiarogodności określenia obciążeń badanego zespołu [24], liczebność próbki badanej na stanowisku [2]. Agamirov i Reicher przyjmowali wartość współczynnika uwzględniającego ewentualną nieadekwatność programu prób do rzeczywistych warunków pracy zespołu równą 1,0 [2], a Ignatowicz z zespołem wartość równą 1,5 [14]. Uwzględniając dostępność dla kontroli miejsca zużycia, charakter postępującego zniszczenia oraz prędkość destrukcji Otshu i zespół przyjmowali wartość współczynnika 1,2 [19]. Uwzględniając liczebność próbki, w przypadku jednej próbki badanej na stanowisku przyjmowano wartość współczynnika 5, a dla sześciu próbek wartość 3 [2]. Generalnie wartość współczynnika bezpieczeństwa przejmowana jest od 1 do 5 [14, 24]. Na podstawie wyników badań zużyciowych oraz uwzględniając współczynniki bezpieczeństwa ustala się trwałość normatywną (resurs) [1, 26].

Innym podejściem szacowania trwałości zespołów hydraulicznych wykorzystującym koncepcje bezpiecznej trwałości są badania niezawodności eksploatacyjnej [17, 27]. Strategia ta zakłada użytkowanie zespołu na statku powietrznym do chwili wystąpienia uszkodzenia. W ramach tej strategii wykorzystuje się metody statystyczne oraz komputerowe techniki symulacyjne i programowane badania niezawodności. Strategia ta może być stosowana tylko wówczas, gdy następstwa uszkodzeń nie naruszają zasad bezpieczeństwa pracy i nie zwiększają kosztów eksploatacji zespołów hydraulicznych [27].

Metody szacowania trwałości oparte na współczynnikach bezpieczeństwa nie dają możliwości oceny funkcji rozkładu trwałości zespołu hydraulicznego na etapie projektowania. Dlatego też realizowane są również prace mające na celu zapewnienie efektywnej eksploatacji zespołów hydraulicznych, wykorzystując do tego celu nowoczesne metody diagnostyczne [16, 20]. Głównym kierunkiem tych prac jest opracowywanie metodologii zarządzania prognostycznego i zarządzania stanem technicznym zespołów oparte na łączeniu wielu źródeł informacji z eksploatacji. Do przetwarzania danych z eksploatacji służą nowoczesne techniki śledzenia sieci neuronowych, a także algorytmy automatycznego wnioskowania i algorytmy progresji prawdopodobieństwa awarii [6, 8]. W ten nurt wpisują się również badania tzw. trwałości resztkowej, w których wykorzystuje się modele rozszerzonej techniki filtra Kalmana, przewidywania szeregów czasowych, wielowymiarowego rozkładu danych i rekonstrukcji przestrzeni fazowej [9, 20]. Badany jest również wpływ kontaminacji na trwałość różnych hydraulicznych par precyzyjnych zespołów hydraulicznych [25]. Do prognozowania okresu użytkowania zespołu hydraulicznego wykorzystywana była eksperymentalna metoda pomiaru wrażliwości na zanieczyszczenia oparta o model wrażliwości na zanieczyszczenia. W niektórych pracach identyfikowano czas niezawodnej pracy odnawialnego obiektu technicznego poprzez zastosowanie trzech kryteriów, w których użyto następujących statystyk: zmodyfikowanej statystyki Kołmogorowa-Smirnowa (MK-S), statystyki średniego odchylenia bezwzględnego dystrybuanty hipotetycznej od empirycznej oraz statystyki obliczanej na podstawie zlogarytmowanej funkcji wiarygodności [3, 21]. Wartości tych statystyk posłużyły do rangowania jedenastu rozkładów prawdopodobieństwa uszkodzeń. Wykazano, że na podstawie zagregowanego kryterium uwzględniającego trzy statystyki zgodności dopasowania zwiększa się wiarygodność estymacji rozkładu czasu pracy do uszkodzenia, unikając tym samym błędów jakie można popełnić uzależniając się tylko od jednej z nich.

Agamirov i Vestyak oraz Blancke z zespołem wykazali, że w napędach hydraulicznych występuje silne skorelowanie parametrów określających ich stan zdatności z czasem ich użytkowania [1, 7]. Można więc przewidywać chwilę wystąpienia zmiany stanu technicznego zespołu napędu hydraulicznego, pod warunkiem periodycznej kontroli tego stanu [1, 24]. Wykorzystując tę własność autorzy niniejszego artykułu zaproponowali aprioryczno-predykcyjną metodę szacowania trwałości.

Metoda przedstawiana w niniejszym artykule opiera się na obserwacji wybranego parametru kontrolnego zespołu napędu hydraulicznego w czasie jego użytkowania. Kontrola ta ma na celu wykrycie we właściwym czasie stanu przed awaryjnego (dopuszczalnego). Novum metody jest wykorzystanie do wykrycia stanu przed awaryjnego zasady tolerancji uprzedzających na wybrany parametr kontrolny. Tolerancje uprzedzające stanowią zbiór wartości wybranego parametru kontrolnego zawarty miedzy poziomami granicznym i przed awaryjnym (dopuszczalnym). Periodyczna kontrola stanu technicznego zespołu hydraulicznego za pomocą wybranych parametrów kontrolnych umożliwia przewidywanie chwili wystąpienia stanu granicznego zespołu napędu hydraulicznego. Ilościowe charakterystyki zużyciowe zespołów hydraulicznych zmieniają się w czasie a ich wpływ na stan techniczny lotniczego napędu hydraulicznego ma charakter losowy. Warunkiem implementacji metody jest znajomość poziomu granicznego parametru kontrolnego zespołu napędu hydraulicznego. Wartość graniczna parametru kontrolnego zespołu napędu hydraulicznego jest określana na etapie jego konstruowania i projektowania. Wynika ona z warunków konstrukcyjnych hydraulicznych par precyzyjnych (par nurnikowych, rozdzielczych, regulacyjnych) i funkcjonalnych całego zespołu hydraulicznego. Zależy w dużej mierze od zastosowanych materiałów oraz rozwiązania konstrukcyjnego hydraulicznych par precyzyjnych i konfrontowana jest głównie z procesami destrukcji tych par, w wyniku ich eksploatacji. Wartość graniczna najważniejszych parametrów zespołu hydraulicznego podawana jest przez producenta w jego dokumentacji technicznej i stanowi kryterium odniesienia w czasie eksploatacji.

Przedstawiana metoda oparta jest na kontroli poziomu wartości dopuszczalnej (stan przed awaryjny) wybranego parametru kontrolnego i określeniu związku tego parametru z okresowością jego sprawdzeń, przy zapewnieniu zadanego poziomu nieuszkadzalności (wyznaczona a priori niezawodność zespołu hydraulicznego). Dopuszczalny poziom parametru kontrolnego jest to taka jego wartość, przy której parametr ten zmierzony w chwili t_1 nie osiągnie do momentu t_2 poziomu granicznego z prawdopodobieństwem $p(t) \geq p_w$, gdzie p_w jest założonym poziomem prawdopodobieństwa bezawaryjnej pracy zespołu w czasie $\Delta\tau = t_2 - t_1$. Jeżeli wartość dowolnego parametru kontrolnego zespołu hydraulicznego η przekroczy wartość dopuszczalną η_{dop}, ale nie przekroczy wartości granicznej η_{gr}, tj. $\eta_{dop} \leq \eta \leq \eta_{gr}$, to uważa się, że napęd hydrauliczny znajdują się w stanie przed awaryjnym. Osiągnięcie przez parametr kontrolny poziomu dopuszczalnego związane jest ze zmianą częstotliwości kontroli tj. $\Delta\tau = t_2 - t_1$. Wielkość tolerancji uprzedzających $\Delta\eta = \eta_{gr} - \eta_{dop}$ jest związana z częstotliwością kontroli $\Delta\tau = t_2 - t_1$ w taki sposób, aby realizacja procesu zmiany parametru decydującego o stanie technicznym zespołu hydraulicznego, po przecięciu poziomu dopuszczalnego η_{dop} przy przepracowanym czasie $t_1 \leq \tau \leq t_2$ nie przecięła do chwili t_2 poziomu η_{gr} z prawdopodobieństwem $p(t) \geq p_w$. Osiągnięcie przez którykolwiek parametr kontrolny wartości dopuszczalnej umożliwia identyfikację zespołów, które wkrótce mogą osiągnąć stan graniczny. Osiągnięcie przez którykolwiek parametr kontrolny η poziomu granicznego η_{gr} tj. $\eta \geq \eta_{gr}$ oznacza kres trwałości zespołu hydraulicznego tj. konieczność zaprzestania jego użytkowania. Należy tu dodać, że w przypadku zespołów odnawialnych można poddać go procedurze remontowej.

2. Opis procesu zmiany parametru kontrolnego zespołu hydraulicznego

W poniższym artykule przyjęto następujące oznaczenia:

$\eta(t)$ – funkcja losowa parametru kontrolnego,
η_{dop} – dopuszczalna wartość parametru kontrolnego w losowym momencie czasu T_{dop},
η_{gr} – graniczna wartość parametru kontrolnego,
T_1 – czas osiągnięcia przez parametr kontrolny poziomu dopuszczalnego,
T_2 – czas przeprowadzania sprawdzenia stanu technicznego po przekroczeniu poziomu dopuszczalnego (zakres trwałości resztkowej),
x – losowy czas przecięcia przez funkcję losową parametru kontrolnego dopuszczalnego η_{dop} lub granicznego η_{gr}.

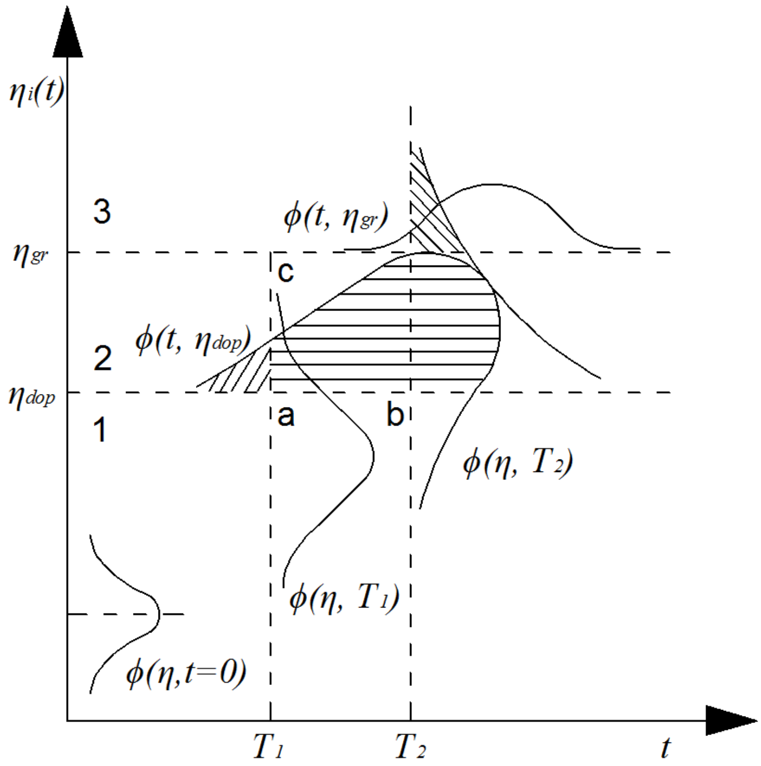
Do opisu metody szacowania trwałości napędu hydraulicznego przyjęto następujące założenia:

1) Zmiany wartości parametru kontrolnego zespołów hydraulicznych przebiegają nieprzerwanie w czasie i zachodzą w wyniku procesów zużycia precyzyjnych par tribologicznych tych zespołów.
2) Zmiana wybranego parametru kontrolnego zespołu napędu hydraulicznego η jest procesem losowym $\eta(t)$ przebiegającym pod oddziaływaniem szerokiego widma czynników eksploatacyjnych.
3) Z badań stanowiskowych lub eksploatacyjnych pozyskano dane umożliwiające formalny opis procesu losowego.
4) Na etapie projektowania określono wartość poziomu granicznego η_{gr} wybranych parametrów kontrolnych zespołu napędu hydraulicznego $\eta(t)$. Wartość graniczna parametru kontrolnego nie ulega zmianie w czasie całego życia zespołu hydraulicznego i jest nieprzekraczalnym kryterium odniesienia.

Aby można było oszacować trwałość zespołu napędu hydraulicznego należy dysponować konkretną postacią rozkładu zmiennej losowej, w postaci funkcji gęstości prawdopodobieństwa.

Na rys. 1 przedstawiono zmiany jednowymiarowej funkcji gęstości rozkładu $\phi(\eta, t)$ losowego parametru kontrolnego i funkcji gęstości rozkładu $f(\eta_{dop}, t)$ przecięcia granicy pola trwałości resztkowej. Przebiegi zmiany funkcji gęstości dzielą trwałość napędu na trzy obszary:

1) obszar w którym zespół hydrauliczny jest w stanie pełnej sprawności,
2) obszar przed awaryjny, w którym występuje ścisły związek wartości tolerancji resztkowej kontrolowanego parametru z okresowością sprawdzeń, przy zapewnieniu zadanego poziomu nieuszkadzalności (wyznaczona a priori niezawodność zespołu hydraulicznego),
3) obszar graniczny, czyli obszar w którym zespół hydrauliczny jest w stanie niezdatności do pracy.



Rys. 1. Charakterystyka trwałości zespołu napędu hydraulicznego dla przypadku procesu losowego $\eta(t)$ zmiany parametru kontrolnego tego zespołu [Źródło: Opracowanie własne]

Z rys. 1 wynika, że dla wykrycia - we właściwym czasie - stanu przed awaryjnego (dopuszczalnego) należy określić związek okresowości sprawdzeń $\Delta\tau = T_2 - T_1$ i tolerancji uprzedzających (trwałości resztkowej) $\Delta\eta = \eta_{gr} - \eta_{dop}$ na kontrolowany parametr, przy zapewnieniu zadanego poziomu nieuszkadzalności (wyznaczona a priori niezawodność zespołu hydraulicznego). Moment sprawdzenia powinien być wybrany w ten sposób, żeby $\eta_{dop} < \eta(T) < \eta_{gr}$.

Dla poziomu parametru kontrolnego η_{dop} mamy $x \leq T_1$ wtedy i tylko wtedy, gdy $\eta > \eta_{dop}$ a dla poziomu η_{gr} mamy $x \leq T_2$ wtedy i tylko wtedy, gdy $\eta > \eta_{gr}$. Stąd do przecięcia zdarzeń na poziomie η_{dop} mamy $\{x \leq T_1\} \cap \{x \leq T_2\} = \{x \leq T_1\}$ wtedy i tylko wtedy, gdy dla czasu T_2 mamy $\{\eta > \eta_{dop}\} \cap \{\eta > \eta_{gr}\} = \{\eta > \eta_{gr}\}$. Możemy więc zapisać, że:

$$P\{x \leq T_1\}_{\eta_{dop}} = P\{\eta > \eta_{gr}\}_{T_2},$$

co oznacza, że prawdopodobieństwo $P\{x \leq T_1\}$ przy poziomie dopuszczalnym parametru kontrolnego η_{dop} jest równe prawdopodobieństwu $P\{\eta > \eta_{gr}\}$ w czasie T_2 sprawdzenia stanu technicznego po przekroczeniu poziomu dopuszczalnego. Stąd otrzymujemy:

$$\int_0^{T_1} f(x/\eta_{dop})dx = \int_{\eta_{gr}}^{\infty} \phi(\eta/T_2,)\, d\eta \qquad (1)$$

gdzie: $f(x/\eta_{dop})$ – warunkowa funkcja gęstości rozkładu wielkości losowej czasu *x* pod warunkiem, że parametr kontrolny przyjął wartość η_{dop};

$\phi(\eta/T_2)$ – warunkowa funkcja gęstości rozkładu wielkości losowej $\eta(t)$ pod warunkiem, że czas pracy osiągnął czas T_2 sprawdzenia stanu technicznego po przekroczeniu poziomu dopuszczalnego.

Podobnie jak równanie (1) wyprowadza się równanie dla poziomu dopuszczalnego η_{dop} w czasie T_2:

$$\int_0^{T_2} f(x/\eta_{dop})dx = \int_{\eta_{dop}}^{\infty} \phi(\eta/T_2)\, d\eta \qquad (2)$$

Porównując równanie (1) z równaniem (2) otrzymamy:

$$\int_{T_1}^{T_2} f(t/\eta_{dop})dt = \int_{\eta_{dop}}^{\eta_{gr}} \phi(\eta/T_2)\, d\eta \qquad (3)$$

Z równania (3) wynika, że dla monotonicznego procesu losowego $\eta(t)$ z danym czasem T_1 i znanej wartości poziomu granicznego η_{gr} wyznaczyć można kolejny termin sprawdzenia stanu technicznego T_2 i wartość poziomu dopuszczalnego η_{dop} w tym czasie. Następstwem wynikającym z zapisu równania (3) jest następujące równanie:

$$\int_{T_1}^{T_2} f\left(t/\eta_{gr}\right)dt = \int_{0}^{T_1} f\left(t/\eta_{dop}\right)dt \tag{4}$$

Z powyższego równania wynika, że zmiana wartości wybranego parametru kontrolnego, po przecięciu poziomu dopuszczalnego η_{dop} przy przepracowanym czasie $t_1 \le \tau < t_2$ nie przetnie do czasu t_2 poziomu η_{gr}. Wszystkie trajektorie procesu losowego parametru kontrolnego przechodząc z obszaru ab (patrz rys. 1) do obszaru bc powodują zmianę częstotliwości sprawdzeń zespołu hydraulicznego.

Zmiany wartości wybranych parametrów kontrolnych zespołu hydraulicznego przebiegają nieprzerwanie w czasie i przejście zespołu hydraulicznych z jednego stanu do drugiego zachodzi w wyniku procesów zużycia precyzyjnych par tribologicznych tych zespołów. Ze względu na fakt, że wystąpienie uszkodzenia elementu zespołu hydraulicznego jest spowodowane przypadkowymi zmianami intensywności procesu zużycia można założyć liniowy przebieg procesu zużycia. Pozwala to opisać proces zużycia precyzyjnych par tribologicznych zespołu hydraulicznego rozkładem normalnym.

Dla rozkładu normalnego wartość oczekiwana $m_\eta(t)$ i odchylenie standardowe $\sigma_\eta(t)$ są aproksymowane zależnościami liniowymi:

$$\begin{aligned} m_\eta(t) &= m_a + m_b t \\ \sigma_\eta(t) &= \sigma_a + \sigma_b t \end{aligned} \tag{5}$$

Stałe współczynniki m_a i m_b określa się wzorami:

$$\begin{aligned} m_a &= \frac{t_{i+1}\, m_\eta(t_i) - t_i\, m_\eta(t_{i+1})}{t_{i+1} - t_i} \\ m_b &= \frac{m_\eta(t_{i+1}) - m_\eta(t_i)}{t_{i+1} - t_i} \end{aligned} \tag{5a}$$

Współczynniki σ_a i σ_b oblicza się z analogicznych wzorów. Funkcje momentów $m_\eta(t)$ i $\sigma_\eta(t)$ określamy z histogramów rozkładu $\phi(\eta,\ t_2)$ (patrz rys. 2 do 4).

Dla rozkładu normalnego funkcja gęstości rozkładu $\phi(\eta,\ t_2)$ wielkości losowej $\eta(t)$ w czasie t_2 sprawdzenia stanu technicznego ma postać:

$$\phi(\eta/t_2) = \frac{1}{\sqrt{2\pi}\,(\sigma_a + \sigma_b t_2)} exp\left[-\frac{(\eta - m_a - m_b t_2)^2}{2(\sigma_a + \sigma_b t_2)^2}\right] \tag{6}$$

Na podstawie zależności (4) funkcja gęstości rozkładu czasu pierwszego przecięcia poziomu przed awaryjnego (dopuszczalnego) $f\left(\eta_{dop}, t\right)$ ma postać:

$$f\left(t/\eta_{dop}\right) = \frac{1}{\sqrt{2\pi}\,(\sigma_a + \sigma_b t)} exp\left[-\frac{\left(\eta_{dop} - m_a - m_b t\right)^2}{2(\eta_a + \sigma_b t)}\right] \frac{d}{dt}\left(\frac{\eta_{dop} - m_a - m_b t}{\sigma_a + \sigma_b t}\right) \tag{7}$$

Podstawiając wyrażenie (6) i (7) do równania (3) wykonując całkowanie i niezbędne przekształcenia otrzymamy zależności na η_{dop} i $\Delta\eta = \eta_{gr} - \eta_{dop}$ dla rozkładu normalnego parametru kontrolnego:

$$\eta_{dop} = \frac{\eta_{gr}(\sigma_a + \sigma_b T_1) - (m_b \sigma_a - m_a \sigma_b)\tau}{\sigma_a + \sigma_b T_1 + \sigma_b \tau} \tag{8}$$

$$\Delta\eta = \frac{\left[(\eta_{gr} - m_a)\sigma_b + m_b \sigma_a\right]\tau}{\sigma_a + \sigma_b T_1 + \sigma_b \tau} \tag{9}$$

Czas osiągnięcia przez parametr kontrolny poziomu dopuszczalnego T_1 tj. czas pierwszego sprawdzania parametru kontrolnego, określić można z warunku założonego poziomu prawdopodobieństwa bezawaryjnej pracy P_{b_p} zgodnie z następującym wyrażeniem:

$$P\{\eta_{gr} < \eta \leq \infty, t_1\} = \int_{\eta_{gr}}^{\infty} \phi(\eta/t_1) d\eta \leq \delta_{dop} \tag{10}$$

gdzie $\delta_{dop} = 1 - P_{b_p}$ jest dopuszczalnym prawdopodobieństwem uszkodzenia.
Podstawiając funkcję gęstości rozkładu $\phi(\eta,\, t_2)$ tj. zależność (6) do wyrażenia (10) określimy czas osiągnięcia przez parametr kontrolny poziomu dopuszczalnego T_1 tj. moment pierwszego sprawdzania parametru kontrolnego, w następującej postaci:

$$T_1 = \frac{\eta_{gr} - m_a - u_{p_{bp}} \sigma_a}{m_b - u_{p_{bp}} \sigma_a} \;, \tag{11}$$

gdzie: $u_{p_{bp}}$ jest kwantylem rozkładu normalnego odpowiadający prawdopodobieństwu P_{b_p}.

Czas pierwszego sprawdzenia zespołu hydraulicznego jako całości (osiągnięcie przez którykolwiek parametr kontrolny poziomu dopuszczalnego) określimy z warunku:

$$\mathrm{t}_1 = \min(\mathrm{T}_{1\mathrm{v}}, \mathrm{T}_{1\mathrm{p}}, \mathrm{T}_{1\delta}) \;, \tag{12}$$

gdzie T_{1V}, T_{1p}, $T_{1\delta}$ są wybranymi parametrami kontrolnymi zespołu hydraulicznego np. maksymalne ciśnienie tłoczenia, współczynnik sprawności objętościowej itp.

3. Szacowanie trwałości rotacyjnej pompy hydraulicznej typu tłoczkowego

Jako przykład określenia czasu osiągnięcia przez parametr kontrolny $\eta(t)$ poziomu dopuszczalnego (zakres trwałości ograniczonej) i czasu przeprowadzania sprawdzenia stanu technicznego po przekroczeniu poziomu dopuszczalnego (zakres trwałości monitorowanej) oraz poziomu dopuszczalnego η_{dop} parametru kontrolnego $\eta(t)$ posłuży nam rotacyjnymi pompami typu tłoczkowego z tarczą rozdzielczą i regulacją wydatku.

W czasie badań pomp rejestruje się między innymi następujące jej parametry kontrolne: maksymalne ciśnienie tłoczenia p_{tmax}, współczynnik sprawności objętościowej ϑ_{vp} i sumaryczne luzy promieniowe w parach tłoczkowych δ_{pt}. Powyższe parametry traktować będziemy jako wielkości losowe, tj. $\eta_p(t_i)$, $\eta_v(t_i)$, $\eta_\delta(t_i)$.

Dla ustalonych wartości czasu pracy t_i tłoczkowych pomp hydraulicznych wynoszących: 0 h, 500h i 1000h, dla każdej wielkości losowej $\eta_i(t_i)$ określa się empiryczną funkcję gęstości rozkładu $\phi(\eta_i,\ t_i)$, wartość oczekiwaną m_i i średnie odchylenie kwadratowe σ_i. Parametry stochastyczne $\phi(\eta_i,\ t_i)$, m_i i σ_i dla parametrów kontrolnych: maksymalnego ciśnienie tłoczenia, współczynnika sprawności objętościowej pompy i sumarycznego luzu promieniowego w parach tłoczkowych uzyskano z badań laboratoryjnych i sprawdzeń kontrolnych w czasie eksploatacji pomp na samolocie, których wyniki znajdują się w opracowaniach wewnętrznych Instytutu Technicznego Wojsk Lotniczych. Podstawiając wartości parametrów kontrolnych do zależności (5a) a następnie wartości tych współczynników do (5) otrzymamy dla założonego czasu pracy pompy funkcje momentów parametrów tłoczkowej pompy hydraulicznej.

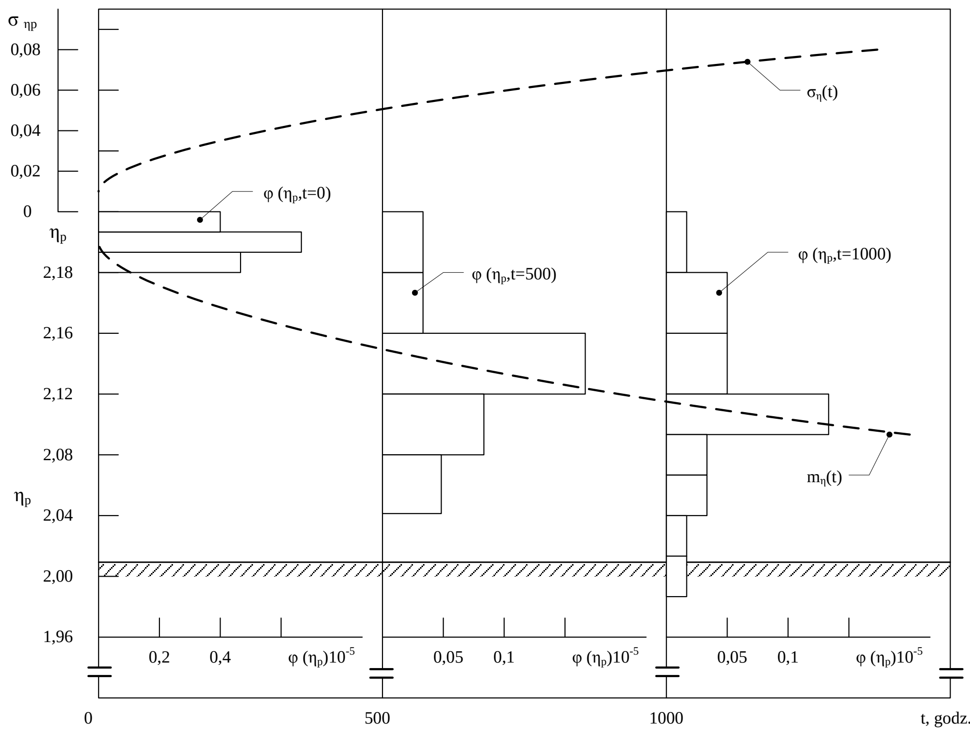
Histogramy rozkładów $\phi(\eta, t)$ i funkcji momentów $m_\eta(t)$, $\sigma_\eta(t)$ dla ciśnienia maksymalnego przedstawiono na rys. 2, współczynnika sprawności objętościowej pompy hydraulicznej na rys. 3 i sumarycznych luzów osiowych w parach tłoczkowych pompy hydraulicznej na rys. 4.

Dla współczynnika sprawności objętościowej ϑ_{vp} funkcje momentów parametrów tłoczkowej pompy hydraulicznej będą:

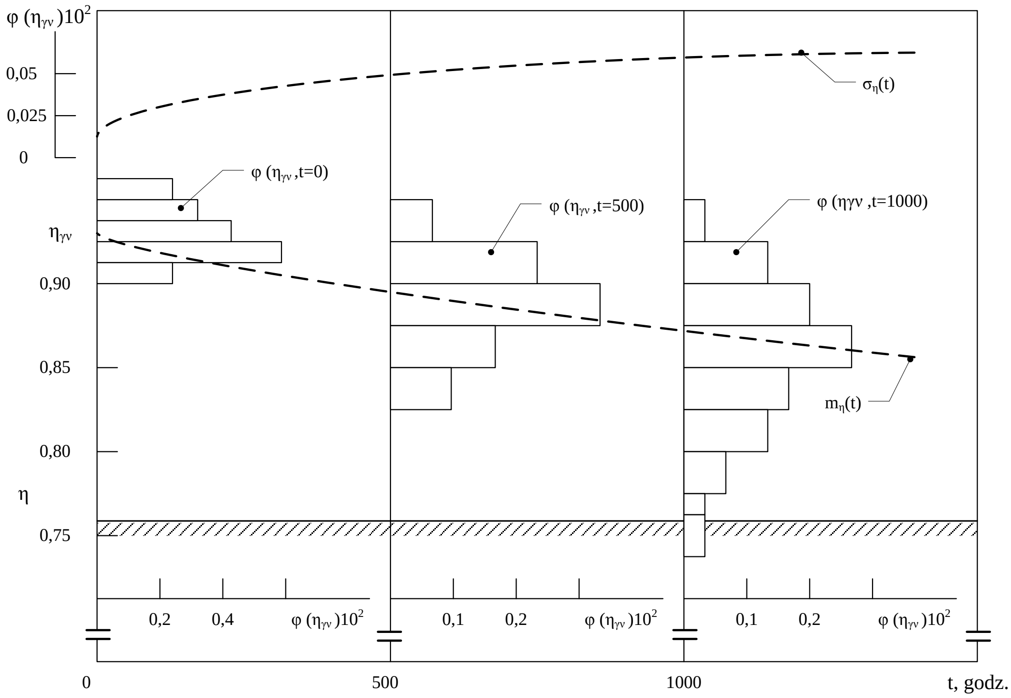
$$m_{\eta_v}(t) = 0{,}942 - 0{,}000065 \cdot t$$
$$\sigma_{\eta_v}(t) = 0{,}024 + 0{,}000015 \cdot t$$

Dla maksymalnego ciśnienia p_{tmax} w [Pa], funkcje momentów parametrów tłoczkowej pompy hydraulicznej będą:

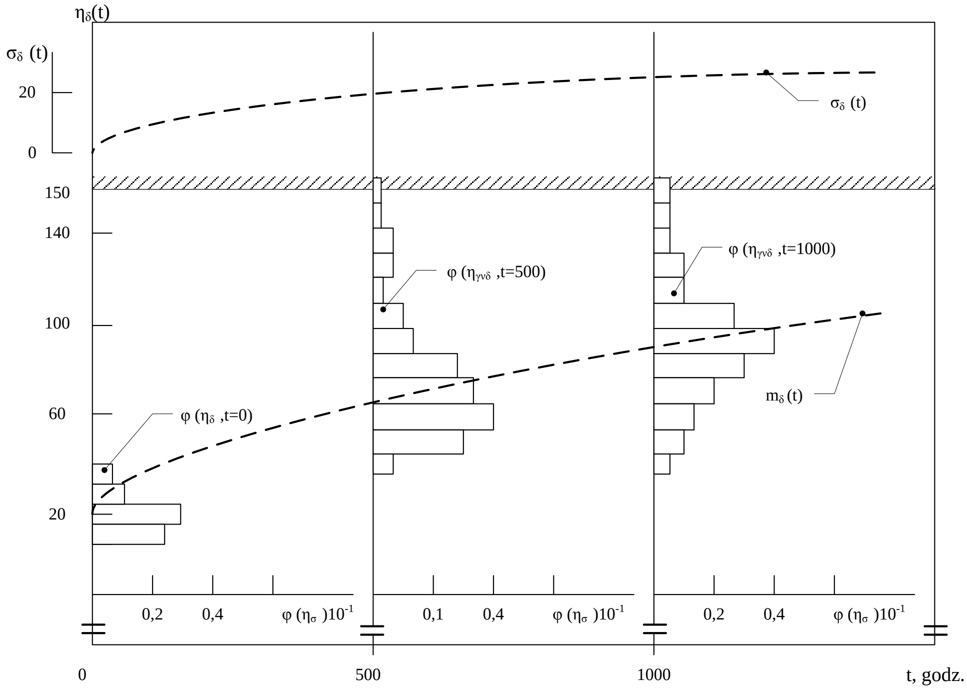
$$m_{\eta_p}(t) = (215{,}6 - 0{,}0031 \cdot t)10^5$$
$$\sigma_{\eta_p}(t) = (3{,}43 + 0{,}00054 \cdot t)10^5$$



Rys. 2. Histogramy rozkładów $\phi(\eta, t)$ i funkcji momentów $m_\eta(t)$, $\sigma_\eta(t)$ dla ciśnienia maksymalnego [Źródło: Opracowanie własne]



Rys. 3. Histogramy rozkładów $\phi(\eta, t)$ i funkcji momentów $m_\eta(t)$, $\sigma_\eta(t)$ dla współczynnika sprawności objętościowej pompy hydraulicznej [Źródło: Opracowanie własne]



Rys. 4. Histogramy rozkładów $\phi(\eta, t)$ i funkcji momentów $m_\eta(t)$, $\sigma_\eta(t)$ dla sumarycznych luzów osiowych w parach tłoczkowych pompy hydraulicznej [Źródło: Opracowanie własne]

Dla sumarycznego luzu promieniowego w parach tłoczkowych δ_{pt} w [μm], funkcje momentów parametrów tłoczkowej pompy hydraulicznej będą:

$$m_{\eta_\delta}(t) = 49{,}34 - 0{,}00973 \cdot t$$
$$\sigma_{\eta_\delta}(t) = 18{,}8 + 0{,}0012 \cdot t$$

Dla pomp hydraulicznych wyznaczono poziom graniczny: współczynnika sprawności objętościowej pompy tj. η_{grv}= 0,75, maksymalnego ciśnienia pompy tj. η_{grp}= 200,9x10^5 Pa, oraz sumarycznego luzu promieniowego w parach tłoczkowych, tj. $\eta_{gr\delta}$= 0,150 µm. Znając poziomy graniczne parametrów kontrolnych określamy wg wzoru (12) czas osiągnięcia przez parametr kontrolny poziomu dopuszczalnego, tj. moment pierwszego sprawdzania parametru kontrolnego.

Dane wyjściowe do określenia funkcji momentów parametrów pompy hydraulicznej oraz zależności $\eta_{dop}(t_i)$ zamieszczono w tablicy 1. Sprawdzenie hipotezy o rozkładzie normalnym $\phi(\eta_i,\, t_r)$ testem zgodności Kołmogorowa wykazało jej zgodność z danymi optymalnymi.

Tablica 1. Dane wyjściowe do określenia momentów parametrów kontrolnych pompy hydraulicznej

Parametry	η_{gri}	m_{ai}	m_{bi}	σ_{ai}	σ_{bi}
Współczynnik sprawności objętościowej pompy	0,750	0,915	-0,000062	0,020	0,000012
Maksymalne ciśnienie tłoczenia [Pa]	$2000{,}9.10^5$	$214{,}9.10^5$	$-0{,}0033.10^5$	$3{,}53.10^5$	$0{,}00059.10^5$
Sumaryczny luzu promieniowy w parach tłoczkowych [µm]	0,150	51,73	0,0397	18,5	0,0012

Czas osiągnięcia przez parametr kontrolny pompy poziomu dopuszczalnego ze względu na jej współczynnik sprawności objętościowej wynosi $t_{1_{vp}}$= 857 godz., ze względu na jej maksymalne ciśnienie $t_{1_{pmax}}$= 1232 godz., ze względu na jej sumaryczne luzy promieniowe w parach tłoczkowych t_{1_δ}= 1326 godz.

Czas osiągnięcia przez parametr kontrolny pompy hydraulicznej poziomu dopuszczalnego określimy z (12):

$$t_1 = min(857, 1232, 1326) = 857\ godzin$$

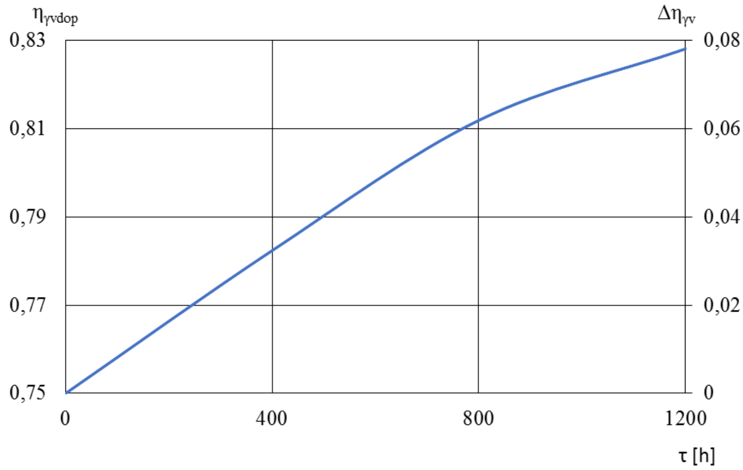
W oparciu o dane wyjściowe przedstawione w tab. 1 za pomocą wzoru (8) określimy zależność poziomu dopuszczalnego parametru kontrolnego η_{dop} od okresowości sprawdzeń dla rozpatrywanych parametrów pompy:

$$\eta_{dopv} = \frac{0{,}0263 + 0{,}00001243 \cdot \tau}{0{,}0268 + 0{,}000012 \cdot \tau},$$

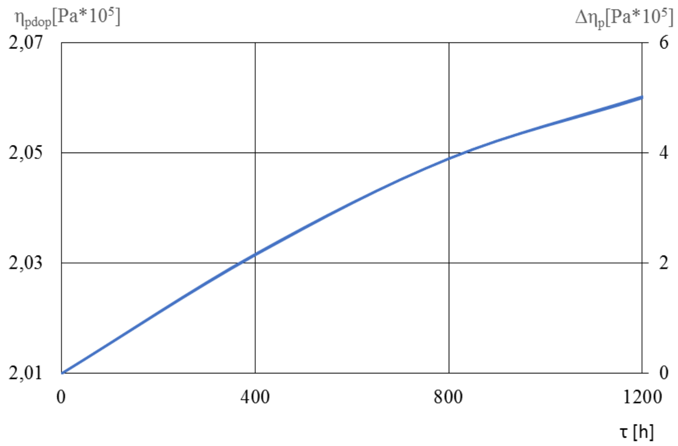
$$\eta_{dopp} = \frac{(801{,}12 + 0{,}1502 \cdot \tau)10^5}{4{,}02 + 0{,}0006 \cdot \tau} \quad [Pa],$$

$$\eta_{dop\delta} = \frac{2879 - 0{,}6879 \cdot \tau}{18{,}95 + 0{,}012 \cdot \tau} \quad [\mu m].$$

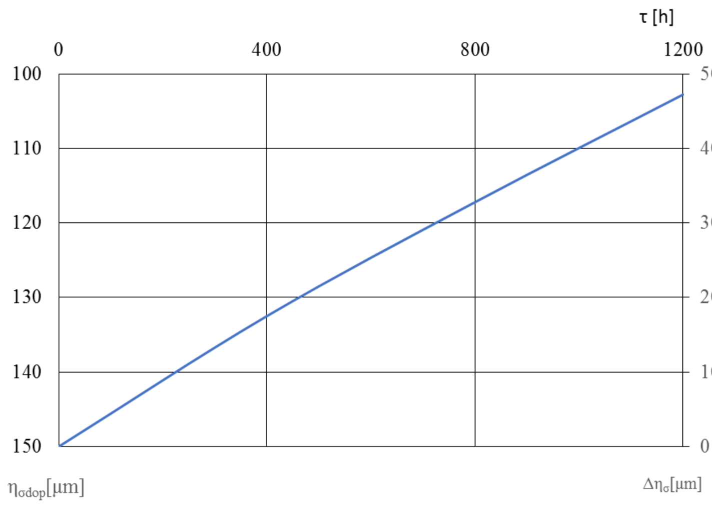
Wartości poziomu dopuszczalnego parametru kontrolnego η_{dop} ze względu na współczynnik sprawności objętościowej pompy przedstawiono na rys. 5, ze względu na maksymalne ciśnienie pompy na rys. 6, ze względu na sumaryczne luzy promieniowe w parach tłoczkowych na rys. 7.



Rys. 5. Zależność poziomu dopuszczalnego η_{dop} i tolerancji resztkowej $\Delta\eta = \eta_{gr} - \eta_{dop}$ od okresowości sprawdzeń τ dla współczynnika sprawności objętościowej pompy [Źródło: Opracowanie własne]



Rys. 6. Zależność poziomu dopuszczalnego η_{dop} i tolerancji resztkowej $\Delta\eta = \eta_{gr} - \eta_{dop}$ od okresowości sprawdzeń τ dla maksymalnego ciśnienia pompy [Źródło: Opracowanie własne]



Rys. 7. Zależność poziomu dopuszczalnego η_{dop} i tolerancji resztkowej $\Delta\eta = \eta_{gr} - \eta_{dop}$ od okresowości sprawdzeń τ dla sumarycznego luzu promieniowego w parach tłoczkowych [Źródło: Opracowanie własne]

Wykresy przedstawione na rys. 5, 6 i 7 wykonano na podstawie obliczeń za pomocą wzoru (8) i (9) dla funkcji i momentów rozkładu $\phi(\eta_i,\ t_i)$, m_i i σ_i parametrów kontrolnych dla czasu pracy t > 500 h. Mają one charakter poglądowy. Prezentują charakter zmiany poziomu dopuszczalnego η_{dop} i tolerancji uprzedzającej $\Delta\eta$ dla wybranego parametru kontrolnego od okresowości sprawdzeń τ. Dla τ = 0 wartość dopuszczalna wybranego parametru kontrolnego osiąga wartość graniczną tego parametru tj. $\eta_{dop} = \eta_{gr}$ a tolerancja uprzedzająca $\Delta\eta$ = 0. Osiągnięty jest kres trwałości zespołu ze względu na konkretny parametr kontrolny. Na podstawie wykresu dotyczącego np. współczynnika sprawności objętościowej pompy możemy określić okresowość sprawdzeń ze względu na ten parametr. Jeśli w czasie kontroli wartość współczynnika sprawności objętościowej będzie wynosiła 0,81 to czas następnego przeglądu będzie wynosił 800 h, natomiast gdyby wartość tego współczynnika wynosiła 0,78 to czas następnego przeglądu będzie wynosił 400 h. Zmiana wartości dopuszczalnej współczynnika sprawności objętościowej w polu tolerancji uprzedzającej powoduje zmianę czasu kontroli (sprawdzenia).

4. Uwagi końcowe

Zaprezentowana metoda szacowania trwałości wykorzystuje występującą w lotniczych napędach hydraulicznych właściwość polegającą na silnym skorelowaniu parametrów określających ich stan zdatności z czasem ich użytkowania. Umożliwia to prognozowanie chwili wystąpienia stanu granicznego zespołu napędu hydraulicznego, pod warunkiem periodycznej kontroli stanu technicznego tego zespołu z wykorzystaniem wybranych parametrów kontrolnych. Kontrola ta ma na celu wykrycie we właściwym czasie stanu przed awaryjnego (dopuszczalnego). W przedstawionej metodzie do wykrycia stanu przed awaryjnego wykorzystuje się tolerancje uprzedzające wybranego parametru kontrolnego.

Przedstawiono związek tolerancji uprzedzających wybranego parametru kontrolnego z okresowością jego sprawdzeń, przy zapewnieniu zadanego poziomu wyznaczonej a priori niezawodności zespołu hydraulicznego. Osiągnięcie przez wybrany parametr kontrolny poziomu przed awaryjnego (dopuszczalnego) związane jest ze zmianą częstotliwości kontroli tj. $\Delta\tau = t_2 - t_1$. Wielkość tolerancji uprzedzających $\Delta\eta = \eta_{gr} - \eta_{dop}$ jest związana z częstotliwością kontroli $\Delta\tau = t_2 - t_1$ w taki sposób, aby realizacja procesu zmiany wybranego parametru kontrolnego decydującego o stanie technicznym zespołu hydraulicznego, po przecięciu poziomu dopuszczalnego η_{dop} przy przepracowanym czasie $t_1 \le \tau < t_2$ nie przecięła do czasu t_2 poziomu η_{gr} z prawdopodobieństwem nie przekraczającym założonego prawdopodobieństwa bezawaryjnej pracy zespołu w czasie $\Delta\tau$. Osiągnięcie przez którykolwiek parametr kontrolny wartości dopuszczalnej umożliwia identyfikację zespołów, które wkrótce mogą osiągnąć stan graniczny. Osiągnięcie przez którykolwiek parametr kontrolny poziomu granicznego η_{gr} tj. $\eta \ge \eta_{gr}$ oznacza konieczność zaprzestania użytkowania zespołu hydraulicznego.

Uszczegółowienie przedstawionej metody polega na związaniu ogólnej zależności wyrażającej trwałość (czas zdatności) z fizycznymi mechanizmami zużywania się zespołów hydraulicznych i degradacji parametrów kontrolowalnych.

Dla zaimplementowania metody niezbędne jest wyznaczenie na etapie projektowania poziomu granicznego η_{gr} parametru kontrolowanego zespołu napędu hydraulicznego $\eta(t)$.

Przedstawiona metoda wykorzystywana jest w pracach mających na celu określanie zasobu pracy napędów hydraulicznych wojskowych statków powietrznych. Metoda umożliwia użytkowanie napędów hydraulicznych według strategii stanu technicznego z kontrolowaniem parametrów.

Bibliografia

1. Agamirov L V, Vestyak V A. Statistical analysis of results of testing aviation products in the conditions of random evaluation. Software products and systems 2017; 1 (30): 124-129.

2. Agamirov L V, Reicher V L. Fatigue life and damage to aircraft structures. Moscow: Ministry of Science and Higher Education of the Russian Federation, 2018.

3. Andrzejczak K, Selech J. Generalised Gamma Distribution in the Corrective Maintenance Prediction of Homogeneous Vehicles. Reliability and Statistics in Transportation and Communication 2019; 519-529. https://doi.org/10.1007/978-3-030-12450-2_50.

4. Ao D, Hu Z, Mahadevan S. Design of validation experiments for life prediction models. Reliability Engineering & System Safety 2017; 165: 22-33.

5. Bansal R K. Fluid Mechanics and Hydraulics Machines (Edition 9th). New Delhi: Laxmi Publications Private Limited (chapter 6), 2011.

6. Bektas O, Jones J A, Sankararaman S, Roychoudhury I, Goebel K. A neural network filtering approach for similarity-based remaining useful life estimation. The International Journal of Advanced Manufacturing Technology 2019; 101: 87-103.

7. Blancke O, Tahan A, Komljenovic D, Amyot N, Claude M L. A holistic multi-failure mode prognosis approach for complex equipment. Reliability Engineering & System Safety 2018; 180: 136-151.

8. Byington C S, Watson M, Edwards D. Data driven neural network methodology to remaining life predications for aircraft actuator components. In IEEE Aerospace Conference Proceedings (IEEE Cat. No.04TH8720) 2004. DOI: 10.1109/AERO. 2004.1368175.

9. Chilton D. Ensuring proper maintenance and repair in projects involving and elevated hight. Fluid Power Journal 2019; 26 (9): 39-42.

10. Ge W, Wang S., Wear condition prediction of hydraulic pump. Journal of Beijing University of Aeronautics and Astronautics 2011; 37: 1410-1414.

11. Gessner T. Analizing a hydraulic system performance. Fluid Power Journal 2018; 25 (9): 12-15.

12. Grinchar N G, Sorokin P A, Karpychev V A, Sergeev K A. Analysis of change in the state of hydraulic drive of machines in operation according to the diagnostic results. Scientia Iranica B 2020; 27(1): 295 -301.

13. Idziaszek Z, Grzesik N. Object characteristics deterioration effect on task reliability - outline method of estimation and prognosis. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2014; 16 (3): 433–440.

14. Ignatowicz S R., Karuskevich M V., Maslak T P., Jutkowicze S S. Resurs and durability of aviation equipment. Kiev: Ministry of Education and Science of Ukraine, 2015.

15. Klarecki K, Hetmańczyk M P, Rabsztyn D. Influence of the selected settings of the controller on the behavior of the hydraulic servo drive. Mechatronics - Ideas for Industrial Application. Advances in Intelligent Systems and Computing 2015: 317: 91-100.

16. Lee J, Wu F, Zhao W, Ghaffari M, Liao L, Siegel D. Prognostics and health management design for rotary machinery systems - Reviews methodology and applications. Mechanical Systems and Signal Processing 2014; 42: 314.

17. Li H S, Chen D N, Yao C Y. Reliability analysis of hydraulic drive system based on evidence theory and Bayesian network. Hydraulic & Pneumatics 2017; 4: 8-14.

18. Modi P N, Seth S M. Hydraulics and Fluid Mechanics Including Hydraulics Machines (19th Edition). Standard Book House, 2013.

19. Ohtsu I, Yasuda Y, Gotom H. Wear and tribological test equipment hydraulic components. Journal of Hydraulic Research 2001; 39 (2): 203-209.

20. Olivares W, Vianna L, Yoneyama T. Predictive Maintenance Optimization for Aircraft Redundant Systems Subjected to Multiple Wear Profiles. Systems Journal IEEE 2018; 12 (2): 1170-1181.

21. Selech J, Andrzejczak K. An Aggregate criterion for selecting A distributionfor times to failure of components of rail vehicles. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22(1): 102-111.

22. Srinivasan R. Hydraulic and Pneumatic Controls. Vijay Nicole Imprints Private Limited. 2/e, 2008.

23. Tomaszek H, Żurek J, Jasztal M. Prognozowanie uszkodzen zagrazajacych bezpieczenstwu lotow statkow powietrznych. Radom: Biblioteka Problemow Eksploatacji, 2008.

24. Wang H W, Teng K N. Residual life prediction for highly reliable products with prior accelerated degradation data. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2016; 18(3): 379-389.

25. Wang X, Lin S, Wang S P. Remaining useful life prediction model based on contaminant sensitivity for aviation hydraulic piston pump. In IEEE/CSAA International Conference on Aircraft Utility Systems (AUS). 2016. DOI: 10.1109/AUS.2016.7748057.

26. Werbińska-Wojciechowska S. Time resource problem in logistics systems dependability modelling. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15(4): 427-433.

27. Zhai Q, Chen P, Hong L, Shen L. A random-effects Wiener degradation model based on accelerated failure time. Reliability Engineering & System Safety 2018; 180: 94-103.